

#3
PATENTS

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

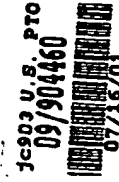
In re application of

Takashi FUJIMORI

Serial No. (unknown)

Filed herewith

HUMAN-MACHINE INTERFACE
SYSTEM MEDIATING HUMAN-COMPUTER
INTERACTION IN COMMUNICATION OF
INFORMATION ON NETWORK



**CLAIM FOR FOREIGN PRIORITY UNDER 35 U.S.C. 119
AND SUBMISSION OF PRIORITY DOCUMENT**

Assistant Commissioner for Patents

Washington, D.C. 20231

Sir:

Attached hereto is a certified copy of applicant's corresponding patent application filed in Japan on July 14, 2000, under No. 2000-215062.

Applicant herewith claims the benefit of the priority filing date of the above-identified application for the above-entitled U.S. application under the provisions of 35 U.S.C. 119:

Respectfully submitted,

YOUNG & THOMPSON

By

Benoit Castel

Benoit Castel
Attorney for Applicant
Registration No. 35,041
Customer No. 00466
745 South 23rd Street
Arlington, VA 22202
Telephone: 703/521-2297

July 16, 2001

OSP-10594 US

日本国特許庁

PATENT OFFICE
JAPANESE GOVERNMENT

US
JCS03 U.S. PRO
09/004460
10/16/01

別紙添付の書類に記載されている事項は下記の出願書類に記載されて
る事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed
this Office.

願年月日
Date of Application:

2000年 7月14日

願番号
Application Number:

特願2000-215062

願人
Applicant(s):

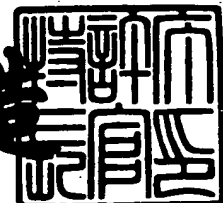
日本電気株式会社

CERTIFIED COPY OF
PRIORITY DOCUMENT

2001年 2月16日

特許庁長官
Commissioner,
Patent Office

及川耕造



出証番号 出証特2001-3008897

【書類名】 特許願

【整理番号】 71310372

【提出日】 平成12年 7月14日

【あて先】 特許庁長官 殿

【国際特許分類】 G06F 3/06

【発明の名称】 ヒューマンマシンインタフェースシステム及びそのプログラムを記録したコンピュータ読取り可能な記録媒体

【請求項の数】 6

【発明者】

 【住所又は居所】 東京都港区芝五丁目7番1号 日本電気株式会社内

 【氏名】 藤森 隆

【特許出願人】

 【識別番号】 000004237

 【氏名又は名称】 日本電気株式会社

【代理人】

 【識別番号】 100108578

 【弁理士】

 【氏名又は名称】 高橋 詔男

【代理人】

 【識別番号】 100064908

 【弁理士】

 【氏名又は名称】 志賀 正武

【選任した代理人】

 【識別番号】 100101465

 【弁理士】

 【氏名又は名称】 青山 正和

【選任した代理人】

 【識別番号】 100108453

 【弁理士】

【氏名又は名称】 村山 靖彦

【手数料の表示】

【予納台帳番号】 008707

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【包括委任状番号】 9709418

【ブルーフの要否】 要

【書類名】 明細書

【発明の名称】 ヒューマンマシンインタフェースシステム及びそのプログラムを記録したコンピュータ読取り可能な記録媒体

【特許請求の範囲】

【請求項1】 ヒューマンマシンインタフェースの機能を分散オブジェクトの形態で構成する各機能要素を、ネットワークのノード上に複数配備し、

各ノードが連携してデータ処理を行うことで、各ヒューマンマシンインタフェースの機能を実現する

ことを特徴とするヒューマンマシンインタフェースシステム。

【請求項2】 前記各ノードが、入出力部であるアプリケーションノードと、アプリケーションノードで入出力されるデータを処理するサービスノードと、両ノードの機能を備える複合ノードのいずれかであることを特徴とする請求項1記載のヒューマンマシンインタフェースシステム。

【請求項3】 前記サービスノード又は複合ノードとして、表現媒体に依存するデータ処理を行う下位サービスノード又は下位複合ノードと、表現媒体に依存しないデータ処理を行う上位サービスノード又は上位複合ノードとが存在し、上位サービスノード又は上位複合ノードが、複数の異なる表現媒体にそれぞれが依存する複数の下位サービスノード又は下位複合ノードによって共通に利用されることを特徴とする請求項2記載のヒューマンマシンインタフェースシステム。

【請求項4】 前記アプリケーションノード又は複合ノードから他の前記サービスノード又は複合ノードに対してサービス開始要求と処理データとを送り、該他の前記サービスノード又は複合ノードにおいて該アプリケーションノード又は複合ノードにおける入力情報又は出力情報の作成処理を行うことを特徴とする請求項2又は3記載のヒューマンマシンインタフェースシステム。

【請求項5】 前記各ノードが、上位から下位に向けて、アプリケーションオブジェクト又はサービスオブジェクトと、分散オブジェクトの上位部分であるプロキシと、分散オブジェクトの下位部分であるオブジェクトトランスポート機構及びリモートクラスリファレンス機構と、ネットワークトランスポート層と、ネットワークインタフェース回路とからなるソフトウェア実行機構を有して構

成されていることを特徴とする請求項 1 ～ 4 のいずれか 1 項に記載のヒューマンマシンインタフェースシステム。

【請求項 6】 請求項 1 ～ 5 のいずれか 1 項に記載のヒューマンマシンインタフェースシステムにおいて実行されるプログラムを記録した計算機読取り可能な記録媒体。

【発明の詳細な説明】

【 0 0 0 1 】

【発明の属する技術分野】

本発明は、人間とコンピュータの間の情報交換を仲介するためのヒューマンマシンインタフェースシステム及びそのプログラムを記録したコンピュータ読取り可能な記録媒体に関する。

【 0 0 0 2 】

【従来の技術】

従来のヒューマンマシンインタフェースシステムの一例としては、組込み用のマイクロプロセッサを中心として当該機器に搭載されたハードウェア及びソフトウェア資源のみを使用して実現されているものがあげられる。図13に示すように、音声による操作と応答の機能を有するこの従来のヒューマンマシシインフェースシステムは、機能構成上は音声信号のデジタルデータ化手段1210と、音声認識のための音声信号の前処理手段1211と、音声認識のためのパターンマッチング手段1212、音声認識のための系列判定手段1213と、音声認識結果に基づいて当該装置の動作を制御する機器制御手段1215と、当該装置の内部状態に基づいて操作者に対して情報の提供を行うメッセージ生成手段1216と、生成されたメッセージを音声波形に変換する音声合成手段1217と、音響信号に変換するデジタルデータアナログ信号変換手段1218と、上記の各機能ブロックの一連の動作の制御を行うシステム制御手段1214とから構成されている。そしてこれらの構成のうち、パターンマッチング手段1212ではパターン要素マッチング処理を行うときに所定の記憶装置に記憶されている認識用パターン辞書1220を参照し、系列判定手段1213では系列判定処理を行うときに所定の記憶装置に記憶されている認識用単語辞書1221を参照し、メッセージ生成手段1216ではメッセージ生成処理を行うときに所定の記憶

装置に記憶されている合成用単語辞書1222を参照し、そして、音声合成手段1217では音声合成処理を行うときに所定の記憶装置に記憶されている合成用パターン辞書1223を参照する。

【 0 0 0 3 】

また回路要素構成上では、同じく図13に示す様にアナログ音声信号入力回路及びA/D変換回路1204が音声信号のデジタルデータ化を行い、信号処理プロセッサ1202が環境雑音の排除及び特徴パラメータの抽出等のデジタル音声データの前処理を行い、同じかまたは別の信号処理プロセッサ1202が特徴パラメータの標準単位パターンとのマッチング処理を行い、同じ信号処理プロセッサまたは別のプロセッサ1202が系列の判定を行う。機器の制御プロセッサ1201が判定された結果に基づいて当該機器の制御を行い、内部状態の情報提供のためのメッセージを生成する。音声認識処理に使用したのと同じのまたは別の信号処理プロセッサ1202はメッセージから音声を合成して、D/A変換回路及びアナログ音声出力回路1203により音声波形として送出する。上記の処理に共通な回路要素として音声信号の蓄積や処理の中間結果、制御のためのプログラム実行に必要な記憶回路が含まれる。こうしたシステムが動作するのに必要な電源回路、タイミング生成回路も含まれている。

【 0 0 0 4 】

しかし、上記の従来技術を複数の機器からなるヒューマンマシンインタフェースに適用する場合には、次のような問題点があった。第1の問題点は、従来の方法を用いた場合、ヒューマンマシンインタフェース機能を実現するためのコストが大きいということである。その理由は、想定している組込み用途のマイクロプロセッサベースのシステムとしては音声を用いたヒューマンマシンインタフェースの機能に消費されるハードウェア及びソフトウェアリソースの比率が相対的に大きくかつ個々の機器がすべて同機能の為のリソースを具備する必要があるためである。また、多くの場合にこうしたヒューマンマシンインタフェースの機能は、当該機器にとってはその主たる目的ではなく、操作性を向上するための手段として採用されている事から費用対効果の評価が低くなされる傾向が強い。

【 0 0 0 5 】

第2の問題点は、搭載可能な機能及び性能が不十分ということである。その理由は、上記のコスト上の理由と製品に設定可能な上限コストにより実際に搭載されるヒューマンマシンインタフェースはその機能や性能が十分高いものを搭載することが困難なためである。そのほかの機能・性能の制限要因としては小型の機器の場合や携帯型の機器の場合、電源容量の制限、放熱容量の制限などから高性能の信号処理プロセッサやマイクロプロセッサ、大容量のメモリを搭載することは事実上困難な点である。

【0006】

第3の問題点は異なる複数の機器の間でヒューマンマシンインタフェースに関わる情報が有効利用されないということである。その理由は、ヒューマンマシンインタフェースの動作パラメータ等の情報を明示的ないしは適応的に設定することによってその操作性が高められるのであるが、個々の機器が独立にそれらを行い連携がなされないことからすべての機器をそのたびごとに設定する必要がある為である。

【0007】

次に他の従来技術について説明する。上記とは別の従来のある種のヒューマンマシンインタフェースシステムの一例が、特開平10-207683号公報に記載されている。図14に示すように、この従来 of ヒューマンマシンインタフェースシステムは、電話網を介して入力された音声の認識とそれに対する応答処理を効果的に行うことを目的としたものであり、私設交換機装置(PBX)1304と、音声応答装置1300と、音声認識合成サーバ1310、リソース管理装置1311と、ローカルネットワーク1308と、電話回線1302と、電話回線加入者線1306とから構成されている。このような構成を有する従来 of 電話応答に適用されたヒューマンマシンシステムはつぎのように動作する。

【0008】

すなわち、交換機1304を介して音声応答装置1300に着呼があると音声応答装置1300はローカルネットワーク1302を介してリソース管理装置1311に対して利用可能な音声認識サーバを問い合わせる。リソース管理装置はその時点で利用可能な音声認識装置を調べ利用可能な場合には音声応答装置1300に対して音声認識合成

サーバ1310が利用可能であることを伝える。音声応答装置1300は音声認識合成サーバ1310に対して音声信号を送送する。音声認識合成サーバ1310は音声認識処理を行い結果を音声応答装置1300に返す。音声応答装置1300はリソース管理装置1311に対し利用可能な音声合成装置を問い合わせる。リソース管理装置はその時点で利用可能な音声合成装置を調べ利用可能な場合には音声応答装置1300に対して音声認識合成サーバ1310が利用可能であることを伝える。音声応答装置1300は、音声合成テキストを音声認識合成サーバ1310に対して送る。音声認識合成サーバ1310は音声の合成を行い結果を音声応答装置1300に伝送する。音声応答装置1300はPBX1304を介して電話回線1302を通じて合成された音声での返答を返す。

【 0 0 0 9 】

しかし、この従来技術には、次のような問題点があった。第1の問題点は、開放型のシステムでの維持管理コストが煩雑となり、それを実行する場合には大きなコストを要するということである。その理由は、プログラミングモデルが通信プロトコルに強く依存しているため、特にネットワークプロトコルの低位階層の構成が変更し難く、システムの拡張性に関してダイナミックな再構成や複数のプロトコルの混在が要求される機能不特定の民生機器ノードによって構成される環境では維持や管理に多くのコストが必要となるためである。図15は、図14に示すシステムにおけるプログラミングモデルの構成を示すブロック図である。音声応答装置1300で動作するアプリケーションプログラム1401と、音声認識合成サーバ1310で動作するサーバプログラム1411との間には、それぞれのプログラムの下位階層に位置するネットワークトランスポート層1405及び1415と、ネットワークインタフェース回路1406及び1416とが存在している。そして、アプリケーションプログラム1401とサーバプログラム1411では、ネットワークトランスポート層1405及び1415の構成に合わせた専用のインタフェースをそれぞれ用いて、データ伝送が実施される。

【 0 0 1 0 】

第2の問題点は、サービスの処理の仕組みがコマンドレスポンスの形態を基本としていることから、アプリケーションプログラムインタフェースの拡張に伴う変更の影響が及ぶ範囲が広いために長期間にわたってシステムを拡張しつづける

ことが難しいということである。その理由は、新たなインタフェース構造を導入する場合にも、それが影響を与えるすべてのノードのソフトウェアに対してプログラムの更新を行う必要があり、それはネットワーク上で動作している可能性のある過去のインタフェースに対してインタオペラビリティを保証する必要があるためである。

【0011】

本発明の有効性を高める背景としてさらに、昨今の機器のネットワーキングコストの低下と普及の進展があげられる。従来に比してヒューマンマシンインタフェース機能をネットワーク上に構成しようとした場合、その実現に要するコストが低下し、また提供できるバンド幅も拡大し、さらにそれに加えて搭載している機器やネットワーク接続を前提としている機器が増える兆しを見せているのである。

【0012】

【発明が解決しようとする課題】

上記のような従来技術における構成と、それらの課題についてまとめると次のようになる。従来の構成としては、①ネットワークを利用せず、ヒューマンマシンインタフェースの機能を自装置内で備えるスタンドアロン構成のものと、②ネットワークの接続はなされているものの、ヒューマンマシンインタフェースの機能は個別に持っていて、共通化についても特定用途システムの中で閉じているものに分けることができる。それらのうち、一方、スタンドアロン構成のものでは、従来の機器におけるヒューマンマシンインタフェースが上記の様に操作対象の機器に完全に組み込まれており、他の機器やシステムとのインタラクションは想定されていない。他方、ネットワークを持つものでは、ネットワークを利用して特定のヒューマンインタフェースの機能を共用している構成である。しかし、上記のような例では音声の認識機能をアプリケーションサーバとして配備している。このような場合では、アプリケーションサービスの単位で分散化をしているために、異なる表現媒体での処理機能の共通化はなされていないことから、比較的低レベルの処理までは対処できるが、ヒューマンインタフェースの統合を行うのには不適當な形態である。

【0013】

上述したように、従来の構成においては、個々の機器が個別にヒューマンマシンインタフェースを持つことによって次の欠点が生じていた。①コスト高、②機能不足、使いにくい、③情報の共通化ができない、④適応能力が小さい、⑤利用範囲が狭い。このような欠点を生じる理由は、①類似の機能を複数の機器が独立にもつこと、②搭載できるリソースが価格、スペースにより強く制限を受けること、③全く独立に設計されており、情報を共有する各層の手段が存在しないこと、④リソース制約とネットワーク未対応であること、⑤特定の用途に対して設計されていて、共通の情報を利用できないことである。

【0014】

本発明は、上記の事情を考慮してなされたものであって、機器へのヒューマンマシンインタフェースの組み込み、実装にあたって、従来の問題点の改善を図ることを目的としたものであって、より具体的には、装置あたりのコストを低減すること、あるいは、機能及び性能を引き上げること、あるいは、ヒューマンマシンインタフェースの操作上の情報を機器間で共通化して異なる機器であっても同一の操作感が得られるようにすること、あるいは、システムの拡張を柔軟に行うことができるようにすること、あるいは、異なる表現媒体のヒューマンマシンインタフェースの上位レベルの情報の処理の共通化を図ることができるようにすることを図ったヒューマンマシンインタフェースシステム及びそのプログラムを記録したコンピュータ読取り可能な記録媒体を提供することを目的とする。

【0015】

【課題を解決するための手段】

上記課題を解決するため、請求項1記載の発明は、ヒューマンマシンインタフェースの機能を分散オブジェクトとして分割された形態で構成する各機能要素を、ネットワークのノード上に複数配備し、各ノードが連携してデータ処理を行うことで、各ヒューマンマシンインタフェースの機能を実現することを特徴とする。請求項2記載の発明は、前記各ノードが、入出力部であるアプリケーションノードと、アプリケーションノードで入出力されるデータを処理するサービスノードと、両ノードの機能を備える複合ノードのいずれかであることを特徴とする。

請求項3記載の発明は、前記サービスノード又は複合ノードとして、表現媒体に依存するデータ処理を行う下位サービスノード又は下位複合ノードと、表現媒体に依存しないデータ処理を行う上位サービスノード又は上位複合ノードとが存在し、上位サービスノード又は上位複合ノードが、複数の異なる表現媒体にそれぞれが依存する複数の下位サービスノード又は下位複合ノードによって共通に利用されることを特徴とする。請求項4記載の発明は、前記アプリケーションノード又は複合ノードから他の前記サービスノード又は複合ノードに対してサービス開始要求と処理データとを送り、該他の前記サービスノード又は複合ノードにおいて該アプリケーションノード又は複合ノードにおける入力情報又は出力情報の作成処理を行うことを特徴とする。請求項5記載の発明は、前記各ノードが、上位から下位に向けて、アプリケーションオブジェクト又はサービスオブジェクトと、分散オブジェクトの上位部分であるプロキシーと、分散オブジェクトの下位部分であるオブジェクトトランスポート機構及びリモートクラスリファレンス機構と、ネットワークトランスポート層と、ネットワークインタフェース回路とからなるソフトウェア実行機構を有して構成されていることを特徴とする。請求項6記載の発明は、請求項1～5のいずれか1項に記載のヒューマンマシンインタフェースシステムにおいて実行されるプログラムを記録した計算機読取り可能な記録媒体である。

【0016】

すなわち、本発明は、(1)ヒューマンマシンインタフェース機能をローカルネットワーク等のネットワークによって機能分散し、また情報共有を可能としている点、(2)ヒューマンマシンインタフェースシステムが分散オブジェクトモデルに基づいている点、(3)ヒューマンマシンインタフェースのバックエンドサービスをその処理階層によって分散化したオブジェクトによって実現し、異なる表現メディアのヒューマンマシンインタフェースの高位階層の処理の統合化と情報の共通化を図っている点を特徴とし、それらによって、(4)ヒューマンマシンインタフェースの実現に要するコストを大幅に低下させることを可能にし、(5)機器毎に搭載する場合に比して部品コスト、ソフトウェアコスト、消費エネルギー、実装面積等の制約を大幅に緩和させ、実現できるヒューマンマシンインタフェー

スの機能及び性能の向上を可能にし、(6)開放型のシステムを長期間にわたって維持拡張を容易にかつ低コストで実現することを可能にしている。

【0017】

【発明の実施の形態】

以下、図面を参照して本発明によるヒューマンマシンインタフェースシステムの実施の形態について説明する。

【0018】

本発明は、有線または無線ネットワークで接続された複数の小規模機器において従来の方法に比して高い性能と低いコスト、高い拡張性をもつヒューマンマシンインタフェース機能を提供するものである。ここで、ヒューマンマシンインタフェースとは、コンピュータ等の機械と、人間との接点において、両者の仲介をする機器及びそれを制御するソフトウェアを指している。図1に示すように、その操作及び状態監視の手段としてのヒューマンマシンインタフェースを保有すべき装置が存在するローカルネットワーク上に、それらを実現するのに必要となる複雑で多くの演算量を要するヒューマンマシンインタフェースの機能要素を含む装置を設け、ヒューマンマシンインタフェースの機能全体を直接操作する機器上とヒューマンマシンインタフェースの各階層の処理機能要素を含む複数の装置上に分散させたオブジェクトの連携によって動作する分散オブジェクトモデルに基づく構成とする。このようにして、機能の分散と共有をネットワーク上で階層化したヒューマンマシンインタフェース処理を分散オブジェクトモデルに基づいて実現することによって、ハードウェア資源及び情報資源の効率的な利用が可能となり、個々のヒューマンマシンインタフェースのコストの低減と性能の向上、更には情報の一括管理が可能となることから保守性が改善するとともにシステムの高い拡張性の提供を可能にする。

【0019】

なお、分散オブジェクトモデルとは、一般に、オブジェクト指向プログラミングモデルに基づいて設計実装されたソフトウェアの機能要素がネットワーク(通信機構)で接続された処理装置(ホスト)に分散して配置された系において、それらが様式化された連携の仕組みを通じて相互に呼び出しや参照をおこなうことによ

って系全体として所期のアプリケーションが実現されるソフトウェアのフレームワークを指す。実用化されているものには、OMG(Object Management Group)のCORBA(Common Object Request Broker Architecture)や、SUN Microsystems (商標)のJava (商標) /RMI(及びjini)、Microsoft (商標)のDCOM(Distributed Common Object Model)などがある。

【 0 0 2 0 】

図1を参照すると、本発明のヒューマンマシンインタフェースシステムの一実施形態は、有線または無線を使用した物理層を用いて機器間の通信路を提供するローカルネットワーク100と、個々の機器の本来の仕事をする手段とそのための指令を与え状態を監視ないし確認するためのヒューマンマシンインタフェース手段を有する装置(アプリケーションノード)101、102、103、105と、ヒューマンマシンインタフェース機能のうち特に複雑で処理に要する演算量やハードウェア資源ないしは大量の情報資源が必要な機能を提供することを目的とする手段を搭載した装置(サービスノード)104と、アプリケーションノードでありかつサービスノードでもある装置(複合ノード)106、107から構成されている。なお、この場合、ノードとは、ネットワークを構成するコンピュータ、端末装置、通信制御装置等の機器またはそれを制御するプログラムを指している。

【 0 0 2 1 】

本実施形態において、アプリケーションノードは、コンピュータ、情報機器、通信機器等の端末装置において、人間に対して直接、音声、映像、画像等の媒体や、機械的動作を用いてデータの入出力を行う機能を提供するネットワークの構成要素であり、サービスノードは、アプリケーションノードに対して各種情報処理機能を提供するネットワークの構成要素である。また、本システムでは分散オブジェクトモデルに基づいて各アプリケーションノードと各サービスノード間のデータ処理が実行されるが、その際、各アプリケーションノードにおけるアプリケーションオブジェクトと各サービスノードにおけるサービスオブジェクトとの間のアクセスを可能とするための分散アプリケーションディレクトリサービス及び分散オブジェクトディレクトリサービスは、図示していないサーバ装置によって提供されているものとする。なお、このような分散オブジェクトモデルに関す

る従来の技術の一例が、特開平10-254701号公報、特開平11-96054号公報等に記述されている。

【 0 0 2 2 】

図2を参照すると、図1に示すアプリケーションノード101,102,103,105等に対応するアプリケーションノード200には、その内部機能として、それぞれが単独または統合された形でCPU (Central Processing Unit)、DSP (Digital Signal Processor)、記憶装置、インタフェース等のハードウェアと、それを用いて動作するソフトウェアプログラムとから構成されている、総合制御手段201、ローカルネットワークインタフェース手段202、ディスプレイ処理手段203、音声信号の入力処理手段204、音声信号の出力処理手段205からなる。ただし、ディスプレイ処理手段203、音声信号の入力処理手段204、音声信号の出力処理手段205の各構成については、必ずしもすべてを各アプリケーションノードに設ける必要はなく、1もしくは2個のみとしたり、または、同一の手段を複数系統備えるようにしてもよい。これらの手段はそれぞれ概略つぎのように動作する。

【 0 0 2 3 】

総合制御手段201は、システム制御手段210が中心となり機器制御手段212に対して機器の目的である動作の制御のマクロな制御（複数の制御手順をまとめて行う制御）を行い、ヒューマンマシンインタフェース(HMI)制御手段211に対してそのマクロな指令と監視を行う。ローカルネットワークインタフェース手段202は、分散オブジェクトモデルに基づいたソフトウェアの実行をサポートするとともにネットワークを介したノード間の通信をするための通信処理をおこなう。内部はOSI(Open System Interconnection)参照モデルにおける物理層と一部のデータリンク層の処理を行うNIC(Network Interface Card)220と、データリンク層の一部、ネットワーク層、トランスポート層を含む狭義のネットワークプロトコル処理を行うネットワークプロトコル処理部221、ソフトウェア（通常プログラム）によって構成されるものであって、分散オブジェクト系の実行基盤となる分散オブジェクトインタフェース222で構成される。

【 0 0 2 4 】

ディスプレイ処理手段203は、表示を行う手段を提供するもので、ディスプレ

イのうち複雑な処理や情報資源へのアクセスが必要であるものがネットワークを介してサービスノードへ送られてそこで処理され、その処理結果を受信したものに対して復号処理を行う復号化処理231と、ディスプレイへの表示操作を行う表示操作(ディスプレイ)230で構成される。音声信号の入力処理手段204は、音声入力を行う手段を提供するもので、音声認識等の複雑な処理や情報資源へのアクセスが必要なものがネットワークを介して送られるためその符号化処理を行う符号化処理241と、信号の入力とデジタル化処理を行う変換部240で構成される。音声信号の出力処理手段205は、音声出力を行う手段を提供するもので、テキストからの音声合成等の複雑な処理や情報資源へのアクセスが必要なものがネットワークを介して送られるためその復号化処理を行う復号化処理251と、アナログ信号への変換と信号の出力を行う変換部250で構成される。

【 0 0 2 5 】

なお、上記構成のうち、復号化処理231、符号化処理241、及び復号化処理251は、それぞれ、HMI制御手段211と、ハードウェアあるいはソフトウェアによる通信手段232, 242, 及び252で接続されている。また、本システムにおいては、ヒューマンマシンインタフェースの各データ処理がネットワークの同一の処理系にならないはその代理によって行われる。したがって、各機器101~103(図1)では、それらの処理系との間でデータを授受する部分の構成、すなわち、上記の例ではヒューマンマシンインタフェース(HMI)制御手段211、ディスプレイ処理手段203、音声信号の入力処理手段204、音声信号の出力処理手段205の各構成について、各機器101~103間で比較的容易に共通化することができる。このように各機器のインタフェース仕様を共通化すれば、ヒューマンマシンインタフェースの操作上の情報を機器間で共通化でき、異なる機器であっても同一の操作感を得られることになる。

【 0 0 2 6 】

図3を参照すると、図1に示すサービスノード104等に対応するサービスノード300の内部機能は、それぞれが単独または統合された形でCPU、DSP、記憶装置、インタフェース等のハードウェアとそれを用いて動作するソフトウェアとから構成されている、総合制御手段301、ローカルネットワークインタフェース手段302、

ディスプレイ処理手段303、音声信号の入力処理手段304、音声信号の出力処理手段305からなる。ただし、ディスプレイ処理手段303、音声信号の入力処理手段304、音声信号の出力処理手段305の各構成については、必ずしもすべてを各サービスノードに設ける必要はなく、1もしくは2個のみとしたり、または、同一の手段を複数系統備えるようにしてもよい。これらの手段はそれぞれ概略つぎのように動作する。

【0027】

総合制御手段301は、システム制御手段310が中心となりヒューマンマシンインタフェース(HMI)制御手段311に対してそのマクロな指令と監視を行う。ローカルネットワークインタフェース手段302は、分散オブジェクトモデルに基づいたソフトウェアの実行をサポートするとともにネットワークを介したノード間の通信をするための通信処理をおこなう。その内部は、物理層と一部のデータリンク層の処理を行うNIC320と、データリンク層の一部、ネットワーク層、トランスポート層を含む狭義のネットワークプロトコル処理を行うネットワークプロトコル処理部321と、分散オブジェクト系の実行基盤となる分散オブジェクトインタフェース322とで構成される。ディスプレイ処理手段303は、表示を行う手段を提供するもので、ディスプレイのうち複雑な処理や情報資源へのアクセスが必要であったものを処理し、ネットワークを介して送りだすための符号化処理331と、ディスプレイイメージを生成するディスプレイイメージ生成330で構成される。音声信号の入力処理手段304は、音声入力を行う手段を提供するもので、音声認識等の複雑な処理や情報資源へのアクセスが必要なものの処理を行うためネットワークを介して送られてくる音声信号の復号化処理を行う復号化処理341と、音声認識処理を行う音声認識処理340で構成される。音声信号の出力処理手段305は、音声出力を行う手段を提供するものでテキストからの音声合成等の複雑な処理や情報資源へのアクセスが必要なものが処理されネットワークを介して送られるための符号化処理を行う符号化処理351と、音声合成処理を行う音声合成処理350で構成される。

【0028】

なお、上記構成のうち、符号化処理331、復号化処理341、及び符号化処理351

は、それぞれ、HMI制御手段311と、ハードウェアあるいはソフトウェアによる通信手段332,342,及び352で接続されている。

【 0 0 2 9 】

図4を参照すると、本発明のヒューマンマシンインタフェースシステムの一実施形態の分散オブジェクトモデルに基づくソフトウェア実行の機構が示される。図4に示す構成401～406のうち、アプリケーションオブジェクト401が図2に示すディスプレイ処理手段203、音声信号の入力処理手段204、音声信号の出力処理手段205に対応する構成であり、ブロック402～406が図2のローカルネットワークインタフェース手段202に対応する構成であり、ブロック412～416が図3のローカルネットワークインタフェース手段302に対応する構成であり、そして、サービスオブジェクト411が図3に示すディスプレイ処理手段303、音声信号の入力処理手段304、音声信号の出力処理手段305に対応する構成である。

【 0 0 3 0 】

アプリケーションオブジェクト401は、サービスオブジェクト411を、アプリケーションオブジェクト401とサービスオブジェクト411のそれぞれに接続された下位のレイアを用いて呼び出し、透過的に実行する。下位に位置するプロキシーであるスタブ402及びスケルトン412はそれぞれのローカルホストにおける呼び出しの代行を行い上記透過的执行を実現する。オブジェクトトランスポート403,413はネットワーク上でオブジェクトの参照のための移動機能を提供する。リモートクラスリファレンス404,414はネットワーク上に分散したクラス参照のための機能を提供する。ネットワーク/トランスポート層処理405,415はそれぞれのレイアの通信処理を行いオープンで拡張性の高い通信基盤を提供する。ネットワークインタフェース回路406,416は物理層とデータリンクその一部の処理を行い電気信号によるネットワーク構成手段を提供する。

【 0 0 3 1 】

図4に示す構成において、プロキシー(スタブ)402及びプロキシー(スケルトン)412は、図2及び図3の分散オブジェクトインタフェース222,322をアプリケーションオブジェクト401及びサービスオブジェクト411の構成に依存する上位の部分と依存しない下位の部分に分けた場合に、上位の部分に対応する構成である。オブ

ジェクトトランスポート403,413と、リモートクラスリファレンス404,414とは、ともに図2及び図3の分散オブジェクトインタフェース222,322のうちのアプリケーションオブジェクト401及びサービスオブジェクト411の構成に依存しない下位の部分に対応する構成である。ネットワーク/トランスポート層処理405,415は、例えばTCP/IP (Transmission Control Protocol/Internet Protocol) などのネットワークプロトコル処理を行うものであって、図2,図3のネットワークプロトコル処理部221,321に対応する構成である。そして、ネットワークインタフェース回路406,416は、図2,図3のNIC220,320に対応する構成である。すなわち、これらの構成のうち、スタブ402及びスケルトン412のみがアプリケーションオブジェクト401あるいはサービスオブジェクト411に依存する構成であり、オブジェクトトランスポート403,413～ネットワークインタフェース回路406,416はアプリケーションオブジェクト401あるいはサービスオブジェクト411に依存しない構成である。

【 0 0 3 2 】

次に、図5及び図6、図7、図8のフローチャートを参照して本実施形態の動作について説明する。まず、サービスを提供する1または複数のオブジェクト(サービスオブジェクト411)はネットワーク100上の他の1または複数のアプリケーション(アプリケーションオブジェクト401)が利用できる様に、その存在をネットワーク上のレジストリに登録する必要がある(図5)。起動されたサービスオブジェクトはネットワーク上に存在するレジストリの探索を行う(図5のステップ501)。次に、検索されたレジストリの判定を行い(ステップ502)。このとき登録の要件を満たすレジストリが見出されなかった場合、レジストリ選択の例外処理(図5のステップ550)を行い登録は行われぬ。登録可能なレジストリがある場合には候補のなかから実際に登録するレジストリの選択を行う(図5のステップ503)。決定されたレジストリに対してサービスオブジェクトの登録を行う(図5のステップ504)。さらに、レジストリへの登録の確認を行い(図5のステップ505)、何らかの異常がある場合には登録例外処理を行って異常終了とする(図5のステップ560)。正常に登録されたことが確認された場合には正常終了となる(図5のステップ507)。

【 0 0 3 3 】

次にアプリケーションがサービスを利用とする場合、アプリケーションはまずネットワーク上でレジストリの探索を行う(図6のステップ601)。探索結果のレジストリの判定を行い目的のサービスの登録の有無を調べる(図6のステップ602)、ネットワークのスコープ内に対象とするレジストリが発見されなかった場合にはレジストリの選択例外処理(図6のステップ650)を行い異常終了する。何らかのレジストリが発見された場合にはレジストリの選択結果の決定を行い(図6のステップ603)その内容を参照する(図6のステップ604)。参照の適否の判定を行い(図6のステップ605)、誤りが発見された場合にはサービス参照例外処理(図6のステップ660)を行い異常終了し、誤りがなかった場合にはロードを行った後正常終了する(図6のステップ606)。

【 0 0 3 4 】

次にサービスの利用の具体例を音声生成のケースを用いて説明する。図7に示すように、処理は左側の列の流れ図の示すアプリケーション(アプリケーションオブジェクト401)側の処理(図7のステップ700)と右側の列の流れ図が示すサービス(サービスオブジェクト411)側の処理(図7のステップ720)とが相互作用を行いつつ進行する。まずアプリケーション側では、音声生成サービスについて図6の処理によってサービスの参照を獲得し(図7のステップ701)、サービスに対して利用開始を伝える(図7のステップ702)。一方サービス側は起動処理(図7のステップ721)の後、上記図5の処理によってサービスの登録がなされる(図7のステップ722)。その後サービス開始要求を待っている(図7のステップ723)ときに、サービス要求(図7のステップ702)があると新たに音声生成のためのスレッドを追加起動し(図7のステップ730)、結果を返す。アプリケーション側はその結果によりサービスが開始できることを確認するかタイムアウトによる打ち切りまで応答を待っている(図7のステップ703)ので、音声生成のためのアーギュメントを設定し(図7のステップ704)、音声の生成を指示(図7のステップ705)を行った後、結果待ちの状態に入る(図7のステップ706)。なおアプリケーション側のホストとしてはこの待機期間に他のプロセスを実行することは可能である。

【 0 0 3 5 】

音声生成の指示を受けたサービス側はアーギュメントとして与えられた生成音

声テキストの解析を行い(図7のステップ731)、音響パラメータを決定し、時系列的なパラメータ列を得る(図7のステップ732)。ここで生成に障害を生じるエラーが発見された場合には例外処理(図7のステップ733)を行う。その後パラメータ列から音声波形データを生成し(図7のステップ734)、形式を整えるために符号化を行いアプリケーション側に実行結果として返す(図7のステップ735)。処理が完了したスレッドは削除される(図7のステップ736)。実行結果を待っていたアプリケーション側では音声信号の復号(図7のステップ707)と音響信号として出力するか(図7のステップ708)、更に別のアプリケーションに引き渡す。

【 0 0 3 6 】

次にサービスの利用の具体例をもうひとつ音声認識のケースを用いて説明する。図8に示すように、処理は左側の列の流れ図の示すアプリケーション側の処理(図8のステップ800)と右側の列の流れ図が示すサービス側の処理(図8のステップ840)とが相互作用を行いつつ進行する。まずアプリケーション側では、音声認識サービスを図6の処理によってサービスの参照を獲得し(図8のステップ801)、サービスに対して利用開始を伝える(図8のステップ802)。一方サービス側は起動処理(図8のステップ841)の後、上記図5の処理によってサービスの登録がされる(図8のステップ842)。その後サービス開始要求を待っている(図8のステップ843)ときに、サービス要求(図8のステップ802)があると、新たに音声認識のためのスレッドを追加起動し(図8のステップ850)、結果を返す。アプリケーション側はその結果によりサービスが開始できることを確認するかタイムアウトによる打ち切りまで応答を待っている(図8のステップ803)ので、音声認識の開始を音響的におおまかに検出するための音声入力有無判定(図8のステップ804)を行い、音声の認識を指示(図8のステップ805)を行う。引き続き所定のフレーム単位(例えば1フレーム単位)で符号化(図8のステップ806)、音声有無判定(図8のステップ807)、サービス側への伝送(図8のステップ808)を反復する。発声の終了の検出がされるか所定の時間を過ぎたことによるタイムアウトを検出すると、音声信号終端の指示をサービス側に伝える(図8のステップ810)。

【 0 0 3 7 】

音声認識の指示を受けたサービス側は音声信号の復号(図8のステップ851)、環

境雑音等の除去、より正確な音声区間の判定処理を行う(図8のステップ852)。引き続き音響的な特徴パラメータの抽出を行い(図8のステップ853)、保持している特徴パラメータの辞書とのマッチングを行いその候補に対するスコアリングを逐次行う(図8のステップ854)。ついで、単語辞書とのマッチングとスコアリングを行い(図8のステップ855)、尤度の高いものを選択する操作を音声の終端の指示があるかまたは音声区間終了検出あるいはタイムアウトが発声するまでつづける(図8のステップ856)。認識結果を符号化し(図8のステップ857)、結果としてアプリケーション側に返す(図8のステップ858)。処理が完了したスレッドは削除される(図8のステップ859)。実行結果を待っていたアプリケーション側では(図8のステップ811)、音声認識結果の復号(図8のステップ812)と結果のさらなる加工ないしは別のアプリケーションへの引き渡しを行う(図8のステップ813)。

【 0 0 3 8 】

以上説明したように本実施の形態によれば、次のような効果を得ることができる。第1の効果は、装置あたりのコストを低減することができることにある。その理由は、ネットワークに接続されている機器の間で利用多重度は民生機器等では一般に十分低いためヒューマンマシンインタフェースの個々のサービスの個数は機器の個数に比べてかなり小さく(例えば10%)取ることができるためである。

【 0 0 3 9 】

第2の効果は、機能及び性能を引き上げることが可能になることにある。その理由は、上記の機器一台あたりのコストを低減できると合わせて、機器の電源容量や放熱容量、筐体形状による制約を回避することができるためである。

【 0 0 4 0 】

第3の効果は、ヒューマンマシンインタフェースの操作上の情報を機器間で共通化でき、異なる機器であっても同一の操作感が得られることにある。その理由は、ヒューマンマシンインタフェースの処理はネットワークの同一の処理系にないしはその代理によって行われるためである。

【 0 0 4 1 】

第4の効果は、システムの拡張を柔軟に行うことができることにある。その理由は、より高い処理性能が必要になった場合にはヒューマンマシンインタフェー

スのサービスの多重度を低減するかあるいはより高性能のハードウェアリソースを有するノードを追加することでそれまでに利用していた環境がそのまま利用できるためである。またこのことによって、導入時の費用を低減することが可能となる。

【0042】

第5の効果は、異なる表現媒体のヒューマンマシンインタフェースの上位レベルの情報の処理の共通化を図ることができることにある。その理由は、例えば音声情報と文字情報の処理に共通するテキストの処理、セマンティクスに基づく処理などの上位の処理を独立したサービスとしてネットワーク上に装備することが可能となるからである。

【0043】

次に、本発明の他の実施形態について図面を参照して詳細に説明する。図9を参照すると、本実施形態は、ローカルネットワーク1000と、アプリケーションノードである装置1、装置2、装置3(それぞれ1001、1002、1003)、音声認識サービスノードである装置4(1004)、センテンスレベルのスコアリングを行うノードである装置5(1005)、文字認識ノードでありアプリケーションノードをかねる複合ノードである装置6(1006)、音声生成サービスノードとアプリケーションノードをかねる装置7(1007)とから構成されている。

【0044】

図9に示される様に装置1(1001)、装置2(1002)、装置3(1003)、装置4(1004)、装置5(1005)、装置6(1006)、装置7(1007)はローカルネットワーク1000で接続されている。これらの装置はそれぞれ概略つぎのような機能を有する。装置1、2、3(1001~1003)は、機器固有のアプリケーションの実行とともに操作に対するヒューマンインタフェースのフロントエンドの機能を提供する。装置4(1004)は、ネットワーク1000に接続された上記装置1、2、3(1001~1003)のヒューマンインタフェースのうち音声認識のバックエンド機能を提供する。装置5(1005)は、ヒューマンインタフェースの機能のうち表現媒体に依存することのない上位階層の照合とそれに基づくスコアリングの機能を提供する。装置6(1006)は、ヒューマンインタフェースの機能のうち文字認識のバックエンド機能を提供すると共に装置

固有のアプリケーションを実行する。装置7(1007)は、ヒューマンインタフェースの機能のうち音声生成のバックエンド機能を提供すると共に装置固有のアプリケーションを実行する。

【 0 0 4 5 】

次に、図10及び図11に示すフローチャートを参照して本実施形態における音声認識の動作について詳細に説明する。図10及び図11は一続きのフローチャートを2つの図に分けて示すものであって、図10の結合子(A)と図11の結合子(A)は接続されていて、図10のスレッドS1及びS2は、図11のスレッドS1及びS2と同じものを示している。図10及び図11に示すように、処理は最も左側の列の流れ図の示すアプリケーション側の処理(図10のステップ1100)と中の列の流れ図が示すサービス側1の処理(図10のステップ1140)及び右側の列の流れ図が示すサービス側2の処理(図10のステップ1160)とが相互作用を行いつつ進行する。この場合、アプリケーション側の処理(図10のステップ1100)が、図9の装置1、2、3(1001~1003)のいずれかで実行されるものであって、サービス側1の処理(図10のステップ1140)が図9の装置4(1004)で実行されるものであって、そして、サービス側2の処理(図10のステップ1160)が図9の装置5(1005)で実行されるものである。

【 0 0 4 6 】

まずアプリケーション側では、音声認識サービスを図6の処理によってサービスの参照を獲得し(図10のステップ1101)、サービス1に対して利用開始を伝える(図10のステップ1102)。一方サービス側1は起動処理(図10のステップ1141)の後、上記図5の処理によってサービスの登録がされる(図10のステップ1142)。その後サービス開始要求を待っている(図10のステップ1143)ときに、サービス要求(図10のステップ1102)があると新たに音声認識のためのスレッドS1を追加起動し(図10のステップ1150)、結果を返す。アプリケーション側はその結果によりサービスが開始できることを確認するかタイムアウトによる打ち切りまで応答を待っている(図10のステップ1103)ので、音声認識の開始を音響的におおまかに検出するための音声入力有無判定(図10のステップ1104)を行い、音声の認識を指示(図10のステップ1105)を行う。引き続き所定のフレーム単位(例えば1フレーム単位)で符号化(図10のステップ1106)、音声有無判定(図10のステップ1107)、サービス側1

への伝送(図10のステップ1108)を反復する。発声の終了の検出がされるか所定の時間を過ぎたことによるタイムアウトを検出すると、音声信号終端の指示をサービス側に伝える(図10のステップ1110)。

【 0 0 4 7 】

音声認識の指示を受けたサービス側1は音声信号の復号(図11のステップ1151)、環境雑音等の除去、より正確な音声区間の判定処理を行う(図11のステップ1152)。引き続き音響的な特徴パラメータの抽出を行い(図11のステップ1153)、保持している特徴パラメータの辞書とのマッチングを行い、その候補に対するスコアリングを逐次行う(図11のステップ1154)。ついで、単語辞書とのマッチングとスコアリングを行い(図11のステップ1155)、尤度の高いものを選択する操作を音声の終端の指示があるかまたは音声区間終了検出あるいはタイムアウトが発声するまでづける(図11のステップ1155)。ここで得られた単語ないしはそれに近いレベルの認識結果をサービス側2のセンテンスレベルのスコアリングサービス(図10のステップ1160)に送られる。サービス側2ではこの時点ですでに起動及びサービス登録(図10のステップ1161,1162)と、サービススレッドS2の追加(図10のステップ1170)が、開始要求の受付(図10のステップ1163)の結果として行われている。

【 0 0 4 8 】

センテンスレベルのスコアリングサービススレッドS2では単語辞書の検索(図11のステップ1171)を行い、シンタクス情報によるスコアリング(図11のステップ1172)、セマンティクス情報に依るスコアリング(図11のステップ1173)を経て総合的なスコアリングを行い(図11のステップ1174)、結果をサービス側1に返す処理(図11のステップ1175)が行われる。センテンスが終了した場合(図11のステップ1176)、処理が完了したスレッドS2は削除される(図11のステップ1177)。サービス側1では発声の終了を検出すると(図11のステップ1156)、認識結果を符号化し(図11のステップ1157)、結果としてアプリケーション側に返す(図11のステップ1158)。処理が完了したスレッドS1は削除される(図11のステップ1159)。実行結果を待っていたアプリケーション側では(図11のステップ1111)、音声認識結果の復号(図11のステップ1112)と結果のさらなる加工ないしは別のアプリケーションへの引き渡しを行う(図11のステップ1113)。

【 0 0 4 9 】

次に、図12を参照して、本発明の他の実施形態について説明する。図12に示すシステムは、音声認識応答、文字テキスト表示等を行うヒューマンマシンインタフェースを有するシステムである。図12に示す実施の形態において、システムの全体のハードウェア上の構成はローカルエリアネットワーク(LAN)10に接続されたアプリケーションノードとして機能するホスト11~13と、サービスノードとして機能するホスト14~19と、更にアプリケーションサービスに関するレジストリを提供するホスト20と、分散オブジェクトに対するレジストリを提供するホスト21とから構成されている。ただし、レジストリノードはアプリケーションノードあるいは各種サービスノードと同一ホスト上で稼動することも可能である。またこれらのアプリケーションノード機能/サービスノード機能とホストは動的に変化することが可能である。すなわち、上記の分散オブジェクトないし分散サービスのエンティティは必ずしも異なるホスト上実行されることを要求しない。また、オブジェクトがネットワーク上で引き渡されて他のホスト上で実行される事もあり得る。さらにアプリケーションの要求するバンド幅と伝送遅延が満たされるならば、サブネットワークを持つネットワークにも適用することができる。

【 0 0 5 0 】

まず、アプリケーションノードについて説明する。図12に示す例では、ホスト11は、システムコントロール11aと、HMIコントロール11bと、アプリケーションサービスインタフェース11cと、ネットワークインタフェース(スタブ)11dと、HMI(音声/表示)フロントエンド11eと、アプリケーション特定インタフェースI/O11fとから構成されている。ホスト12とホスト13も、ホスト11と同様に構成されているものとする。これらの構成によって、ホスト11、ホスト12、及びホスト13は、ヒューマンマシンインタフェースサービスのアプリケーションノードとして、音声によるコマンド入力及び音声応答/ステータスディスプレイ機能を提供する。これらノードには何れも装置の目的に応じた制御や入出力機能(アプリケーション特定インタフェースI/O11f)が備わっている。アプリケーションノードの分散アプリケーションインタフェースとして、アプリケーションサービスインタフェース11cと、ネットワークインタフェース11dが存在する。アプリケーション

ノード側のヒューマンマシンインタフェースの統合を行っているのがヒューマンマシンインタフェース(HMI)コントロール11bである。アプリケーションノード側のヒューマンインタフェースのローカルデバイスへのアクセスと制御それに符号化等の信号変換を行うのがHMIフロントエンド11eである。ここでは音声とディスプレイとなっているが、異なる種類のヒューマンインタフェース媒体を使用する場合にはそのタイプに応じてこのレイアは変更される。そして、システムコントロール11aによって各機能が統合制御される。

【 0 0 5 1 】

次に、アプリケーションサービスと、レジストリについて説明する。図12の例では、上述のアプリケーションノード11～13に対してアプリケーションサービスを提供する複数のサービスノード14～17がネットワーク10上に配置されている。この例では、それらは、文字認識サービスノード14、音声認識サービスノード15、音声合成(音声応答)サービスノード16、ディスプレイ内容合成サービスノード17として構成されている。この場合、ノード(ホスト)14は、文字認識サービスコントロール14aと、ロー・レベル文字認識処理14bと、文字認識データ14cと、ネットワークインタフェース(スタブ/スケルトン)14dとから構成されている。ノード15は、音声認識サービスコントロール15aと、音響的音声認識プロセッシング15bと、音響認識データ15cと、ネットワークインタフェース(スタブ/スケルトン)15dとから構成されている。ノード16は、音声合成サービスコントロール16aと、音響的音声生成処理16bと、音響生成データ16cと、ネットワークインタフェース(スタブ/スケルトン)16dとから構成されている。そして、ノード17は、ディスプレイ内容合成サービスコントロール17aと、表示イメージ生成処理17bと、表示イメージ生成データ17cと、ネットワークインタフェース(スタブ/スケルトン)17dとから構成されている。

【 0 0 5 2 】

更に各ヒューマンインタフェースの上位処理にあたる機能を有するシンタクス処理(構文処理)オブジェクト18a、セマンティック/プラグマティック処理(意味/語用処理)オブジェクト19aが、各ノード18,19によって提供される。それらはネットワークを介して各機能を提供するためにネットワークインタフェース(スタ

ブ)18a,19aをそれぞれ有している。この実施形態では、構文処理オブジェクト18a、セマンティック/プラグマティック処理オブジェクト19aによって提供される機能が、共用可能となっていて、文字認識サービス14a、音声認識サービス15a及び音声合成サービス16aのいずれからでも利用できるようになっている。また、ネットワーク上での分散オブジェクト、分散サービスのロケータである分散アプリケーションレジストリ20aと、分散オブジェクトレジストリ21aとが、ホスト20とホスト21によってそれぞれ提供される。

【 0 0 5 3 】

次に、図12に示すシステムの動作について説明する。

(1)オブジェクト、サービスの登録……………ノード14～19がネットワーク10に接続されると、それぞれ分散アプリケーションレジストリ20aと、分散オブジェクトレジストリ21aに対して登録が行われる。典型的なレジストリとしては、前者ではJava(商標) RMI(Remote Method Invocation) レジストリ、後者ではJini(商標) LookupやUPnP(Universal Plug and Play) SSDP(Simple Service Discovery Protocol)プロキシーなどがそれらにあたる。

【 0 0 5 4 】

(2)HMI処理の実行……………ネットワーク10上のあるアプリケーションノード(ノード11とする)がHMI処理、例えば音声認識を行う場合、アプリケーションノード11はネットワーク10上のアプリケーションサービス15を分散アプリケーションレジストリ (20a)を参照することによって発見し、使用開始手続きを行う。サービス要求と共に符号化された音声情報のデータグラムを送る。音声認識サービスノード15は、ここではサービスに対してローカルに存在する音響的なマッチング処理を行うと共に、ネットワーク10上に装備されている構文処理オブジェクト18a及び、セマンティックス/プラグマティック処理オブジェクト19aを起動して入力音声文章の認識処理を行う。結果がサービスノード15からのレスポンスとしてアプリケーションノード11側に送り返される。音声コマンドの受理とそれを反映した内部処理、更に音声による応答のシーケンスの決定などの最上位の処理はここではアプリケーションノード11側のヒューマンマシンインタフェースコントロール11bの部分によって行われる。

【 0 0 5 5 】

②音声応答………音声の応答を行う場合には同様にネットワーク10上の分散アプリケーションサービスである音声合成サービス16aに処理が依頼される。ここでは、応答すべき音声の音響的な合成を行うと共に、合成する文章の構文、意味、状況に応じた修飾を同じくネットワーク10上に装備されている構文処理オブジェクト18a及び、セマンティックス/プラグマティック処理オブジェクト19aを起動して行うことで高品質の音声応答を可能にする。

【 0 0 5 6 】

③表示の生成………アプリケーションノードにおけるグラフィックス/テキストディスプレイによるダイアログの生成においては、同様にネットワーク10上の分散アプリケーションサービスであるディスプレイ内容合成サービス17aに依頼することで、ローカルにはフォントやグラフィックパターンなどの大量のデータを固定的にかつノード毎に重複して持つ必要が無い上に高品質のディスプレイコンテンツを低いプロセッサ負荷で得ることを可能にする。

【 0 0 5 7 】

④その他の適用例………音声HMIのほかに、カメラのイメージの照合やフォーカス調整など、文字の認識サービスなどの性能の改善、コストの低減などに適用可能である。

【 0 0 5 8 】

図12を参照した実施の形態においても、他の実施の形態と同様に、機器の操作者とのインタフェース機能を実現する手段としてヒューマンマシンインタフェースの機能要素を、分散オブジェクトの形態でネットワーク上に配備し、例えば、音声認識処理と音声応答処理において、それぞれに対応して設けられた音声認識サービス15aと音声応答サービス16aとを用いて下位階層の処理を行い、各処理に対して共通に設けられた構文処理オブジェクト18a及びセマンティックス/プラグマティック処理オブジェクト19aを利用して上位階層のデータ処理を行うようにすることで、各処理階層で必要なハードウェアや演算能力、情報等のリソースの共通化を図り、また個々のノードを固有の処理に特化することによって総コストの低減、高い性能の認識応答能力の提供を図ることができる。また、ネットワー

クに接続された機器すべてにわたってヒューマンインタフェースの仕様の共通化が容易となり、更に認識・応答に関する情報の一元化によって適応結果が共通に反映される。それによってヒューマンマシンインタフェース品質や品位が大幅に向上し、製品の価値が高まるとともに利用者の負担が軽減される。

【 0 0 5 9 】

また、ネットワークに接続された機器すべてにわたって、ヒューマンインタフェースに関わるデータ及びプログラムの共通化が可能となるために、同データ及びプログラムの更新、適応化を一元的に行うことができ、システムの構築、保守、拡張を容易に行うことができる。なお、ネットワーク上のヒューマンマシンインタフェース機能要素は、分散オブジェクトとして分散アプリケーションを構成し、分散アプリケーションは分散アプリケーションレジストリによってアプリケーションサービスとして登録され、アプリケーションノードによって参照される。

【 0 0 6 0 】

すなわち、本実施の形態によれば次のような効果を得ることができる。①ネットワークに接続されたヒューマンインタフェース機能を必要とする機器の個々のハードウェアコストを低減することが可能となる。これは、個々の機器が類似の機能を独立に持つ必要がないためである。②ネットワークに接続された機器のヒューマンインタフェース機能の性能の向上、機能の向上を図ることができる。これは、機能の共通化によって、個別に持つ場合に比べて機能あたりのリソースを大きくすることが可能であり、より高性能のハードウェア、ソフトウェアの実装を可能にするからである。③ネットワークに接続された機器のヒューマンインタフェース機能の構築、保守、拡張を一元的に実施することができるためにそれらに要するコストを低減することができる。これは、ヒューマンインタフェースの性能、品質を向上させるのに欠かせない適応化の結果は一元化されており共通に反映されるので個別に行うのに比べて適応効率がよくまた、それらを保守拡張する場合も1箇所に対して行えば良いからである。④リソースの増強が漸進的にできかつ過去に使用していたリソースも継続して使用することができるために維持コストの低減とシステムのライフタイムの長期化を可能にする。これは、分散オ

プロジェクトアーキテクチャに基づいているので、必要とされる処理負荷に応じてリソースを追加増強できるので過剰な初期コストが発生せず、かつ漸次進歩して性能価格比の向上するハードウェアの市場的利点を活用することができるからである。

【 0 0 6 1 】

なお、本発明の実施の形態の応用例としては、例えば、アプリケーションノードとして携帯端末、PDA(Personal Digital Assistants)端末等の携帯形情報機器を用い、ネットワークを無線ネットワークとし、そして、サービスノードとしてワークステーション、大型計算機等を用い、アプリケーションノードをネットワークに対して動的に接続、切断できるようにするもの等が考えられる。

【 0 0 6 2 】

同様な機能を有するシステムを従来の技術によって構成しようとした場合、従来のスタンドアロン型の構成においては、高性能のHMI機能を実現するために、各携帯端末に、高速のプロセッサ、高速のメモリ、大容量のストレージデバイスが必要となり、コストの面で折り合えない。また、携帯形の装置では電源容量の制限が厳しく高性能のハードウェアを搭載できないし、電力消費によって生じる熱を放出するため機構を搭載することが難しい。また、大規模なハードウェアを搭載するスペースの制限が厳しく、また携帯型機器機器毎にそれらの機構を個別に持った場合、情報の共通化が困難である。特に学習など適応化を行う場合に顕著である。さらに、機器毎に個別に持った場合、アップデート等の保守を個別に行う必要がある。また、分散オブジェクトモデルによらない従来のネットワークによるHMIプログラム実行の際の問題点としては、ネットワーク構造やネットワークプロトコルに対する依存性が強く(環境依存性が強く)、民生機器としてシステムを維持管理することが難しいという問題があり、また、様々の種類の機器がネットワークに接続される可能性があり、それらを機能を維持しつつシステムを拡張するのには複雑すぎて困難が多く、HMI統合の効果が充分発揮できない(拡張性が低い)という問題があった。さらに、表現メディア独立(音、画像等の媒体独立)の言語処理が必要であったため、すなわち、従来は音声入力、音声出力、手書き文字入力等の処理を独立に行っていたため、ネットワーク分散による機能

統合がそのままでは利点を発揮できないという問題もあった。これに対して、本発明では、分散オブジェクトモデルとしてシステムを構築しているので、高性能のHMI機能を分散オブジェクトとすることで、携帯端末に備える必要がなくなり、上記のような従来の問題点を解決することができる。また、例えば、上記各実施の形態で採用したように、各サービスを、メディア依存のレイア(文字認識、音声認識、音声合成等の下位の階層のレイア)とメディア独立のレイア(構文処理、セマンティクス/プラグマティック処理等の上位の階層のレイア)に分け、別の機能単位によって実現されるサービスとすることで、異なるメディアにおいて機能を共通化すること、及び辞書等の情報を共用することが可能となり、従来の問題点が解決される。

【 0 0 6 3 】

なお、本発明の実施の形態は上記のものに限定されることはなく、適宜変更することが可能である。例えば、各HMIサービスを提供する各サービスオブジェクトにおいて、例えば認識処理時に、アプリケーションノードとしての端末装置の操作者の識別情報を用いて、操作者毎に認識処理に学習処理を行うようにし、同一の操作者が異なる端末からアクセスしてきた場合にも、過去の学習データを利用した認識処理を実行すること等の構成の追加を行うことができる。また、本発明の各ノードにおいて実行されるプログラムは、全体として、あるいは所定の部分毎に計算機読取り可能な記録媒体を用いて、あるいは通信回線を用いて、配布することが可能である。

【 0 0 6 4 】

【発明の効果】

本発明によれば、従来に比べ、容易に、装置あたりのコストを低減したり、機能及び性能を引き上げたり、ヒューマンマシンインタフェースの操作上の情報を機器間で共通化して異なる機器であっても同一の操作感が得られるようにしたり、システムの拡張を柔軟に行うことができるようにしたり、あるいは、異なる表現媒体のヒューマンマシンインタフェースの上位レベルの情報の処理の共通化を図ることができる。

【図面の簡単な説明】

【図 1】 本発明によるヒューマンマシンインタフェースシステムの一実施の形態を示すブロック図。

【図 2】 図1のアプリケーションノードの構成例を示すブロック図。

【図 3】 図1のサービスノードの構成例を示すブロック図。

【図 4】 図1に示すシステムのプログラミングモデルの構成を示すブロック図である。

【図 5】 図1のシステムの動作を説明するためのフローチャート。

【図 6】 図1のシステムの動作を説明するためのフローチャート。

【図 7】 図1のシステムの動作を説明するためのフローチャート。

【図 8】 図1のシステムの動作を説明するためのフローチャート。

【図 9】 本発明によるヒューマンマシンインタフェースシステムの他の実施の形態を示すブロック図。

【図 1 0】 図9のシステムの動作を説明するためのフローチャート。

【図 1 1】 図9のシステムの動作を説明するためのフローチャート。

【図 1 2】 本発明によるヒューマンマシンインタフェースシステムの他の実施の形態を示すブロック図。

【図 1 3】 従来のヒューマンマシンインタフェースシステムの構成例を示すブロック図。

【図 1 4】 従来のヒューマンマシンインタフェースシステムの他の構成例を示すブロック図。

【図 1 5】 図14に示すシステムのプログラミングモデルの構成を示すブロック図である。

【符号の説明】

10,100,1000 ローカルエリアネットワーク

11,12,13,101,102,103,105,1001,1002,1003,1005 アプリケーションノード

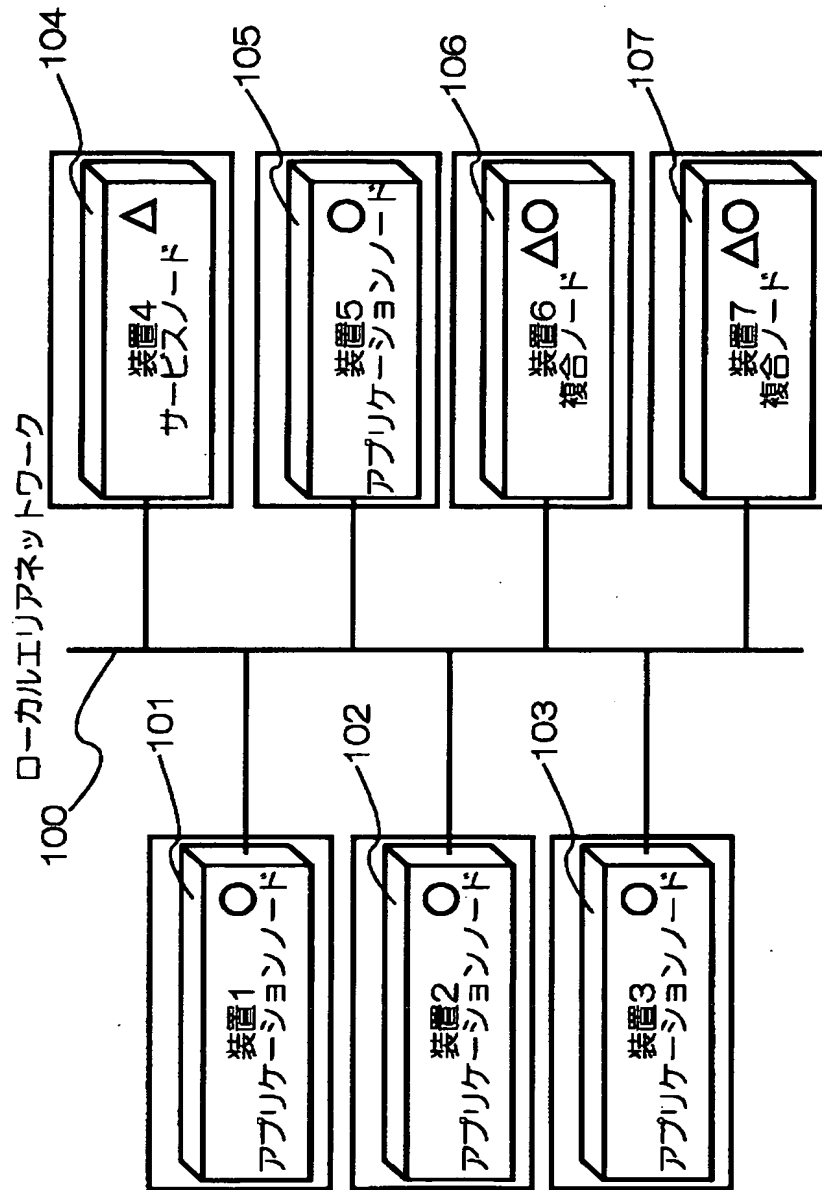
14,15,16,17,18,19,104,1004 サービスノード

106,107,1006,1007 複合ノード

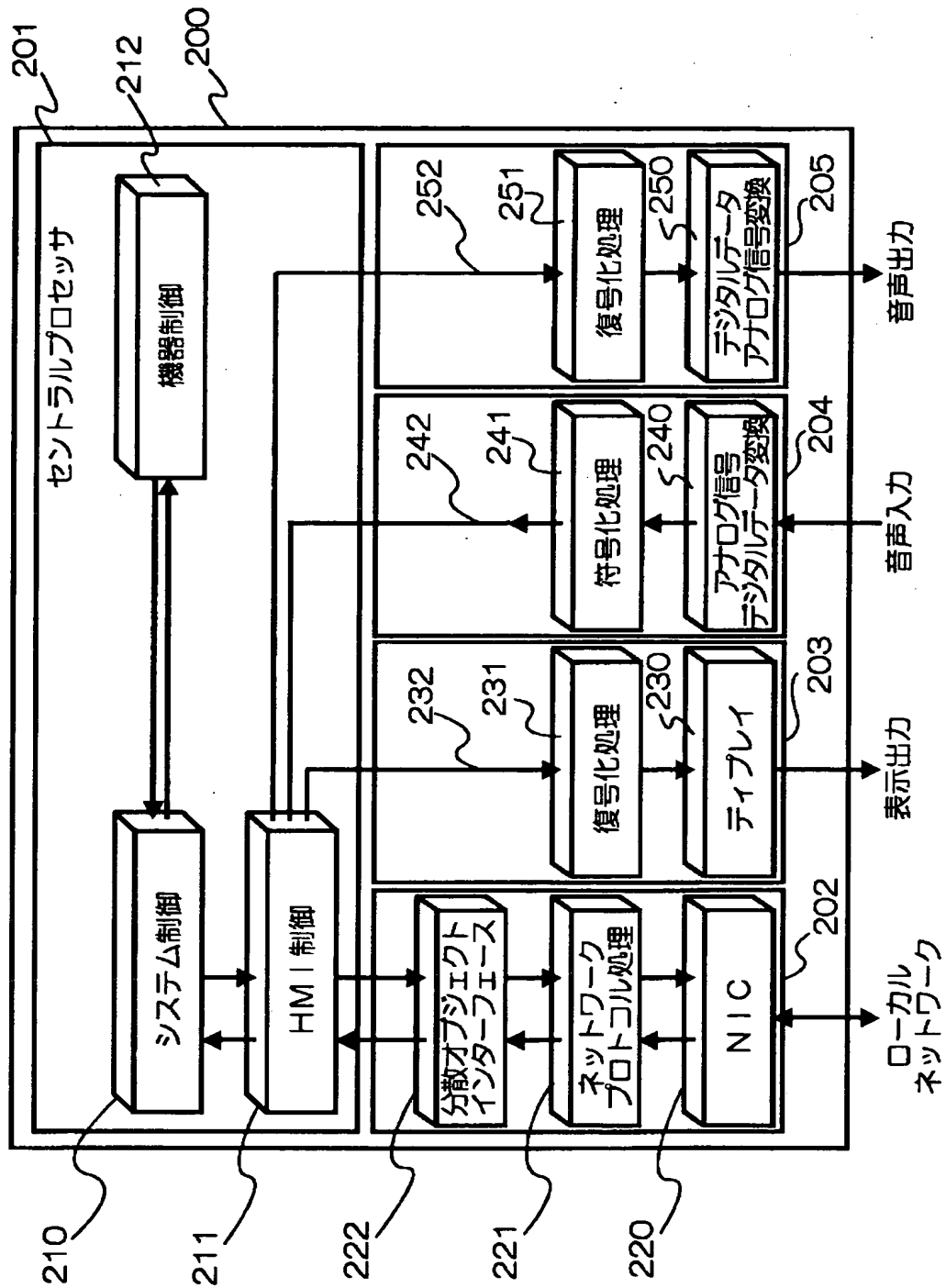
【書類名】

図面

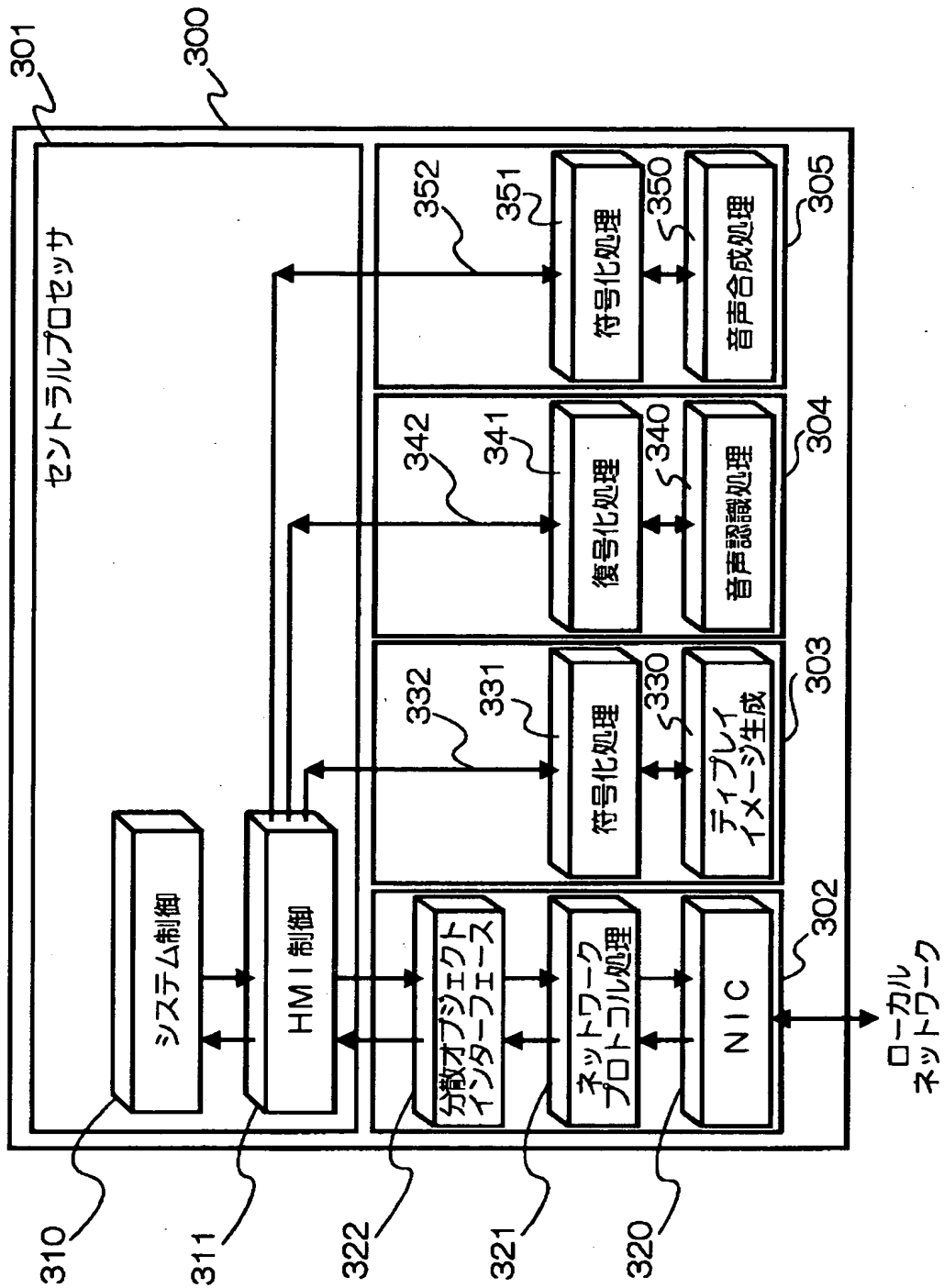
【図 1】



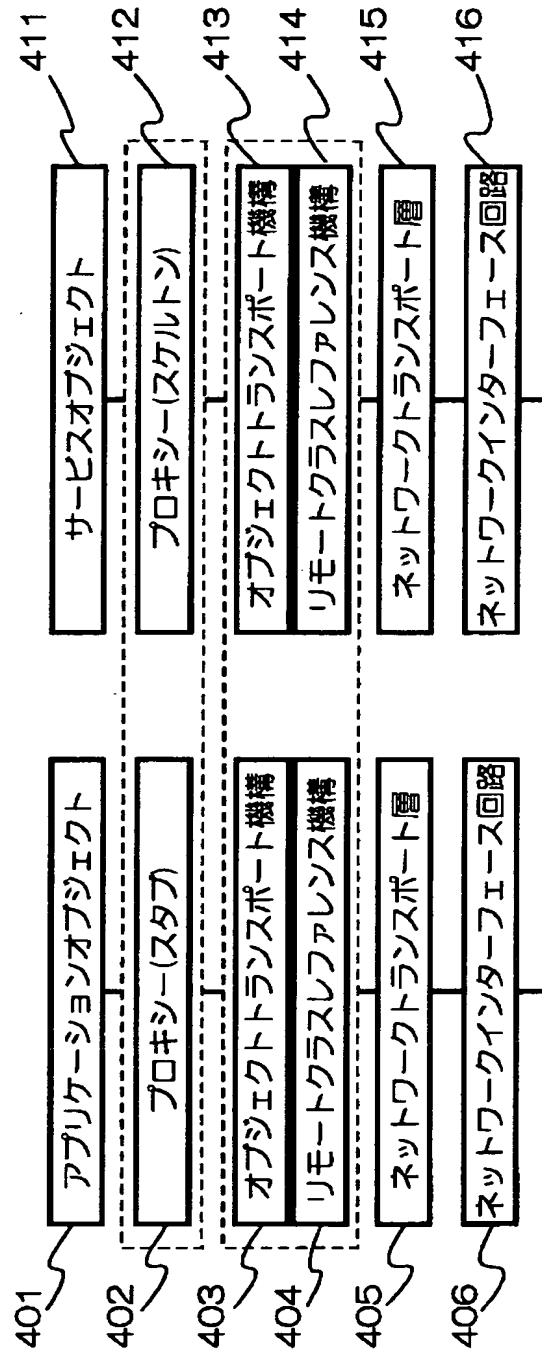
【図 2】



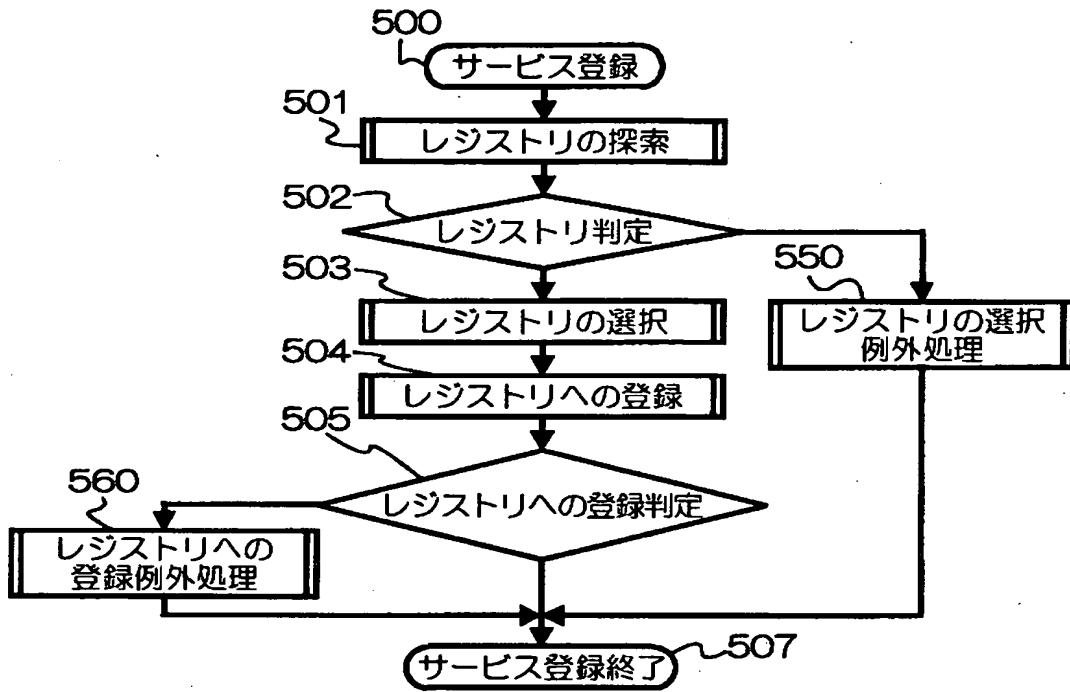
【図 3】



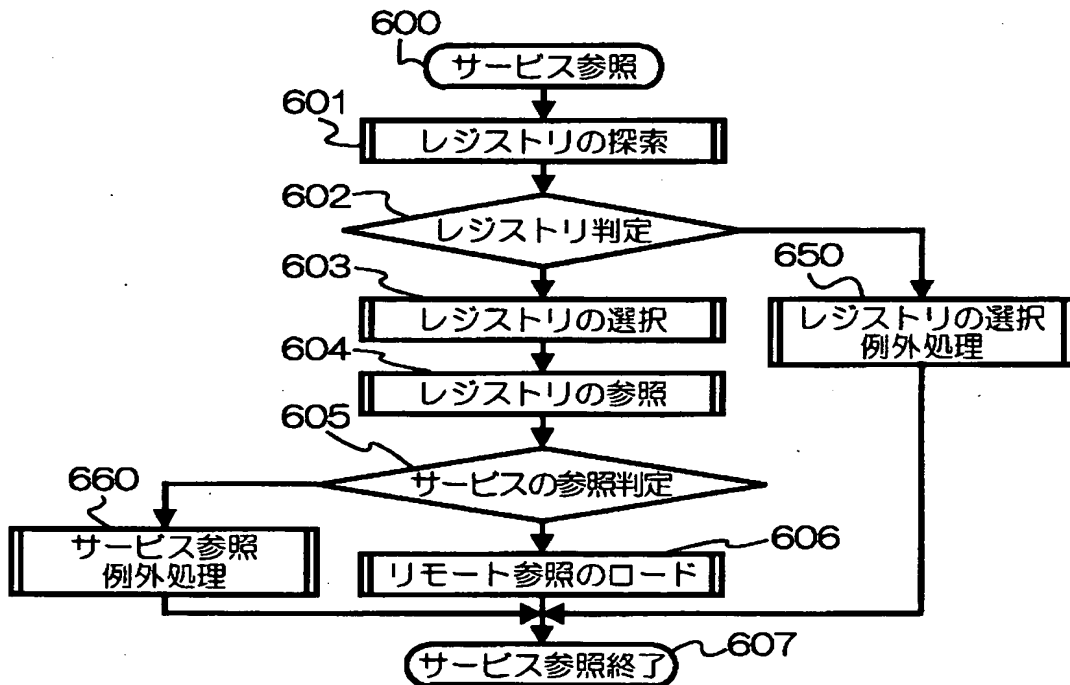
【図4】



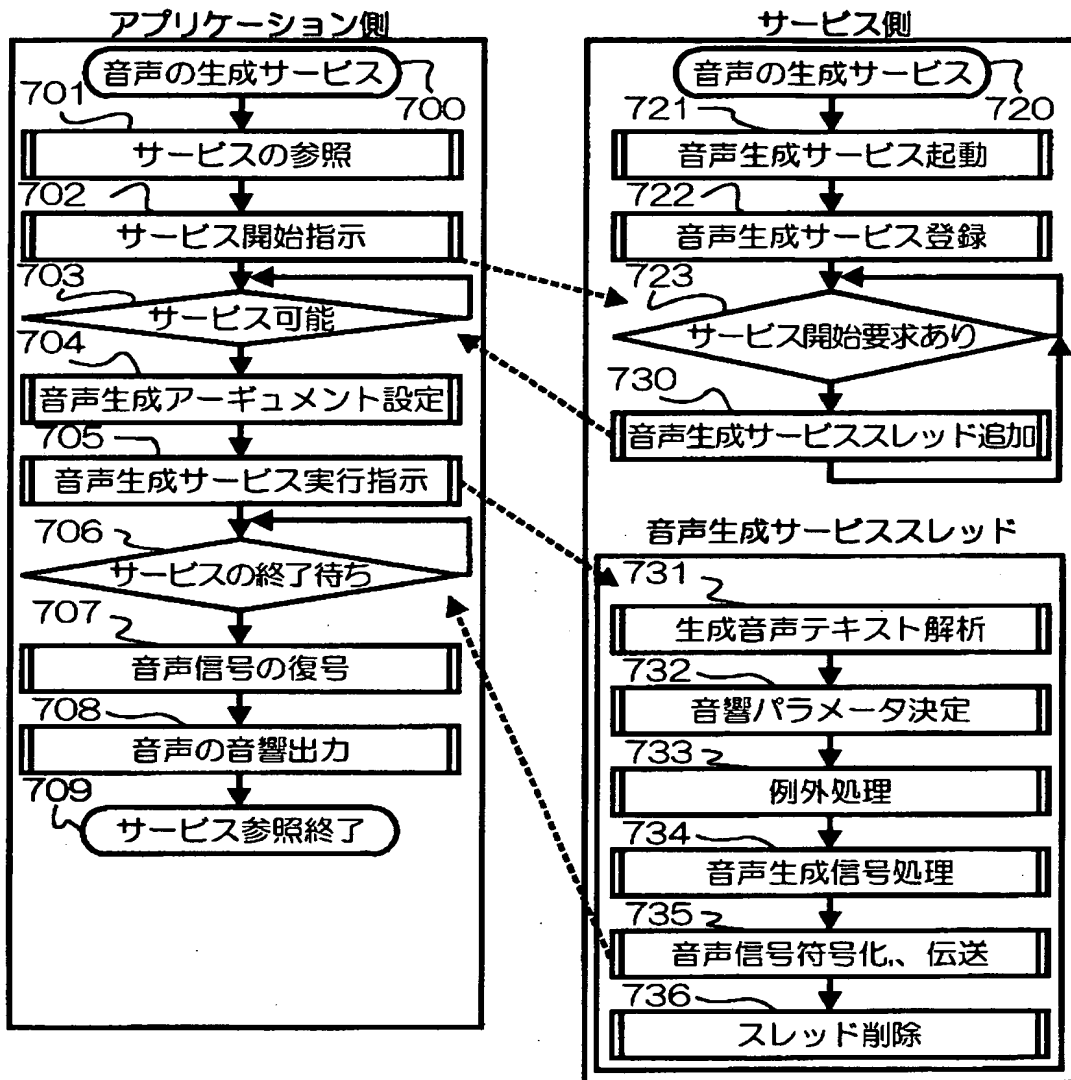
【図 5】



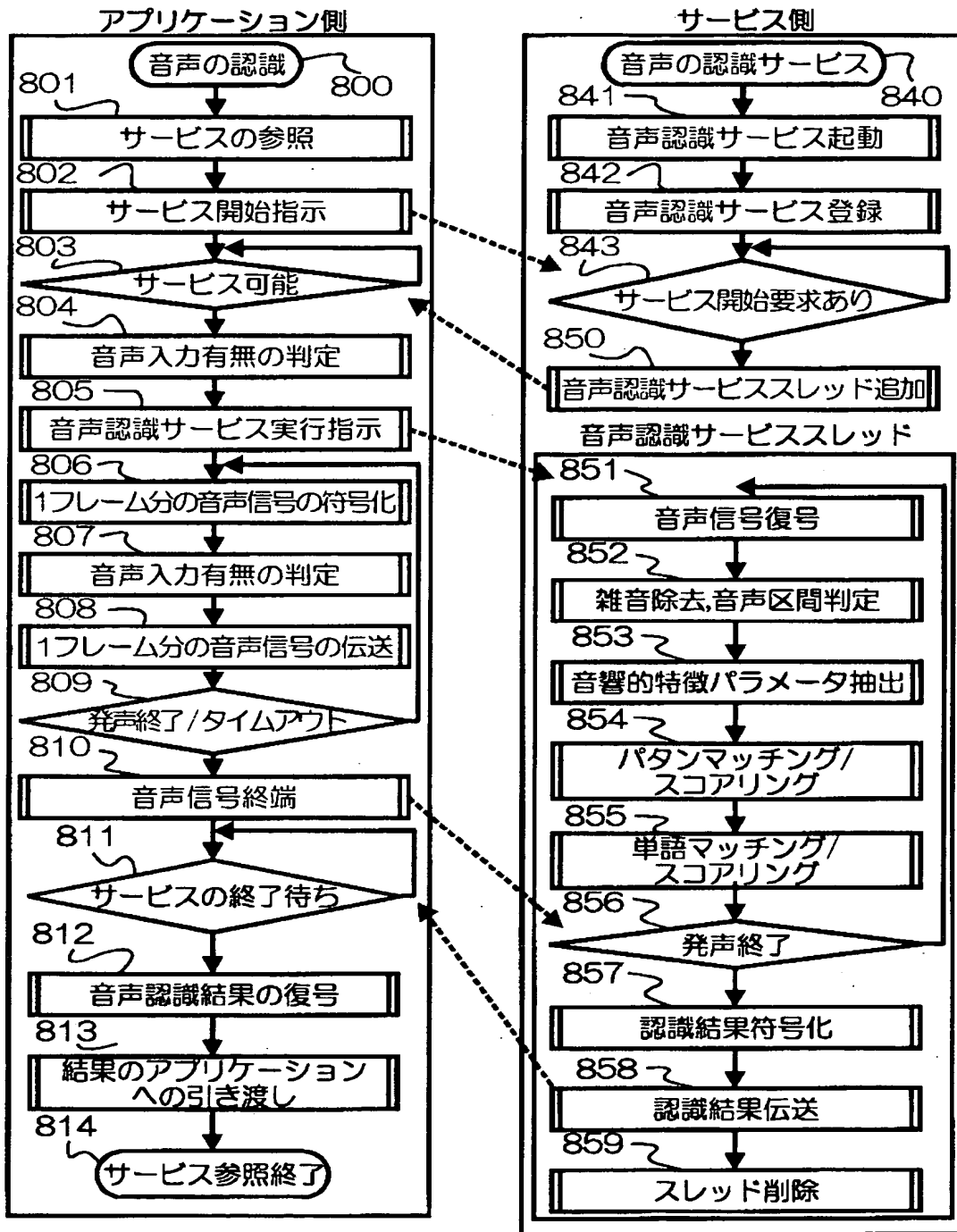
【図 6】



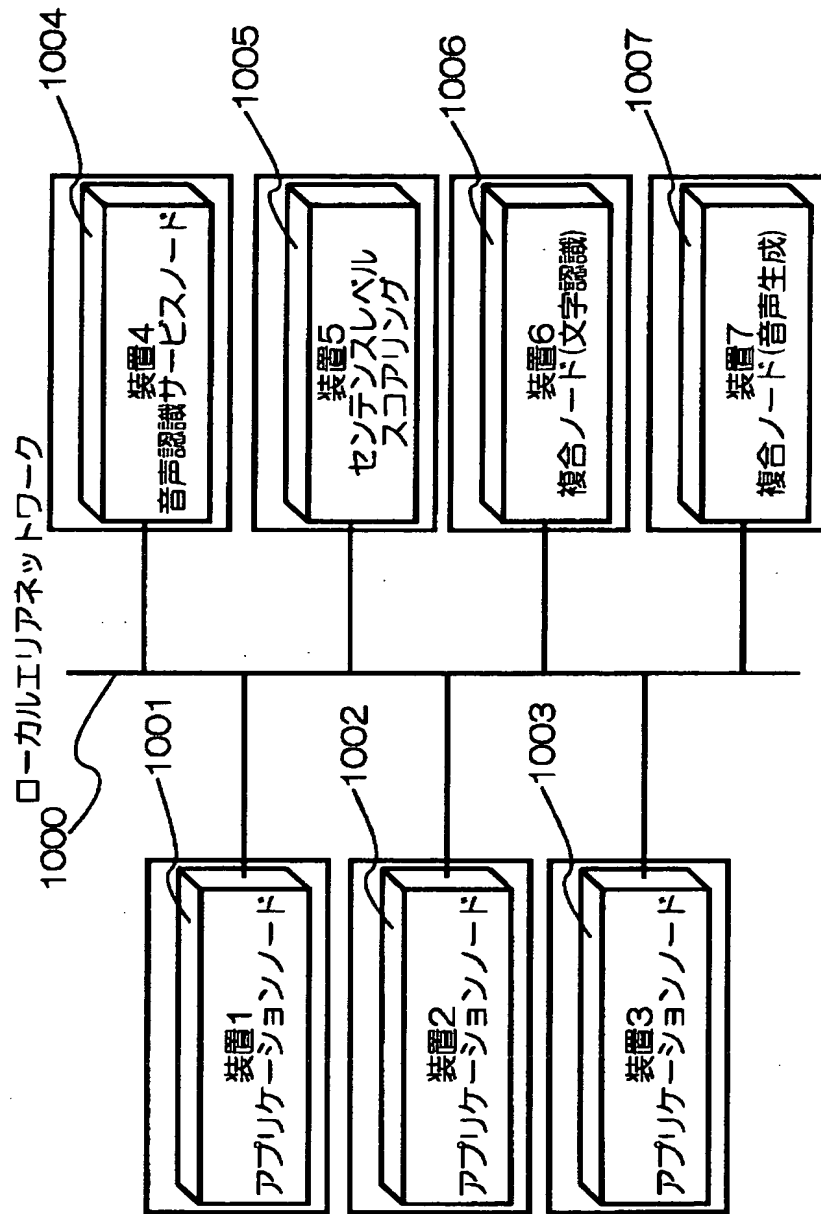
【図 7】



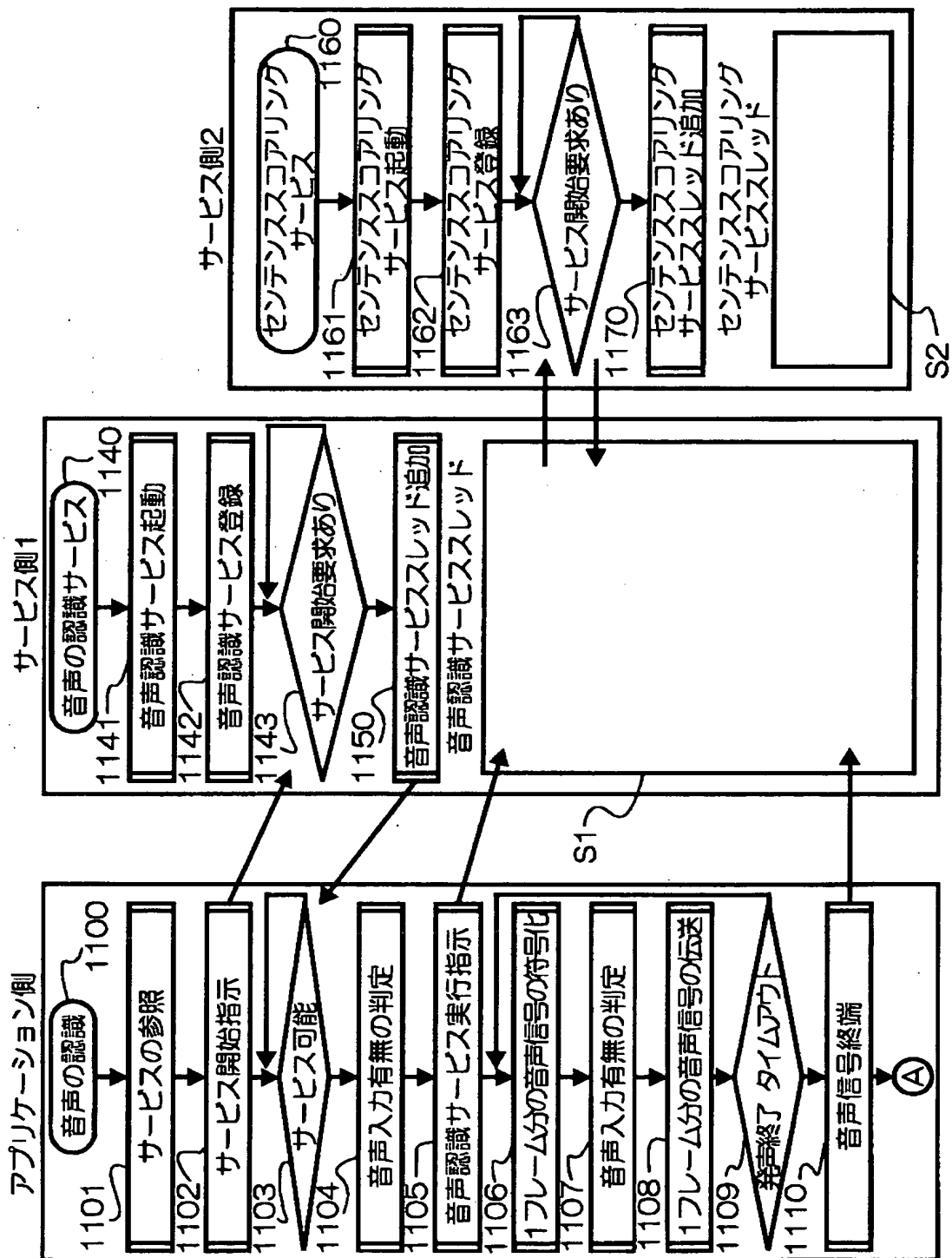
【図 8】



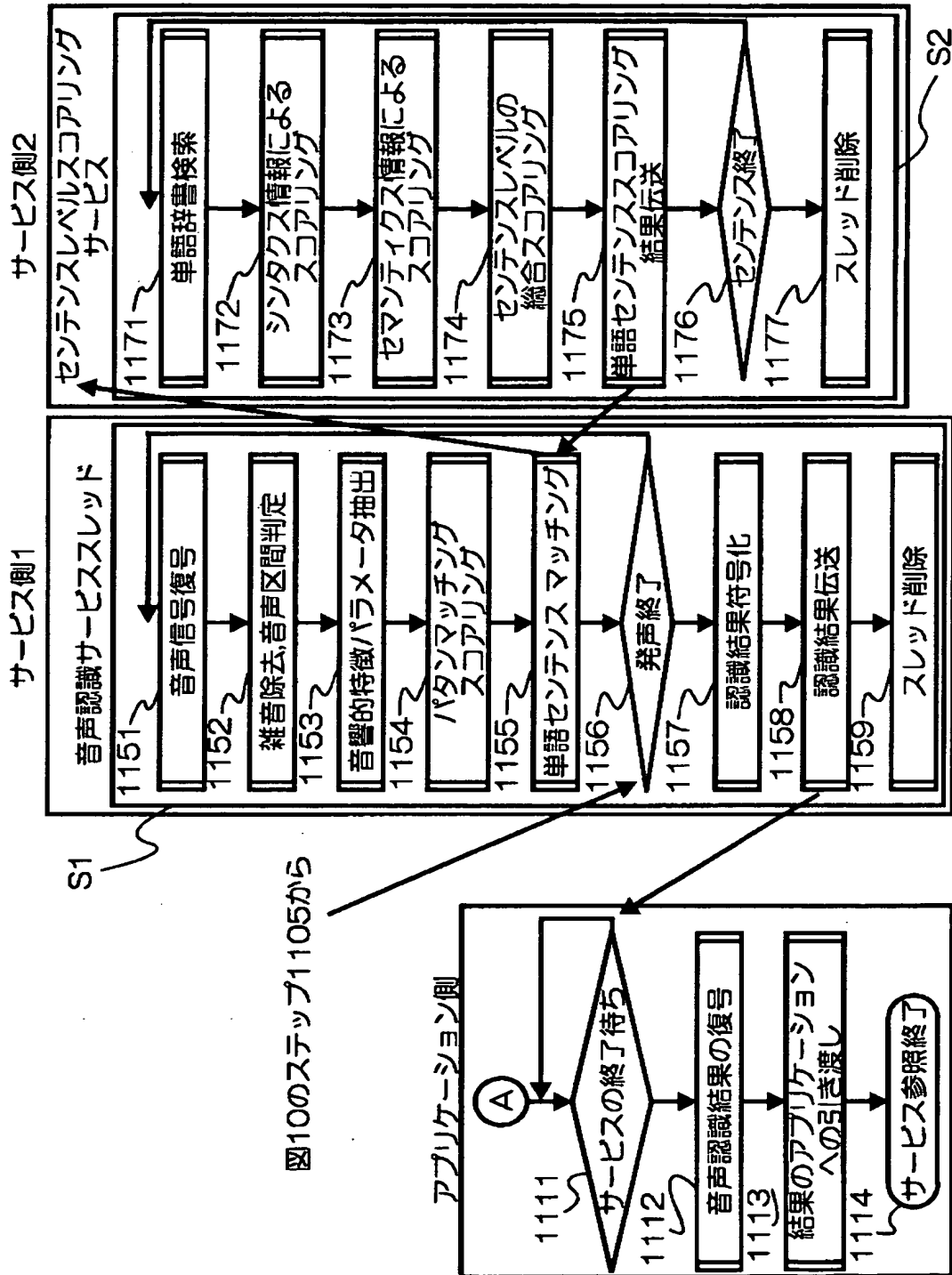
【図9】



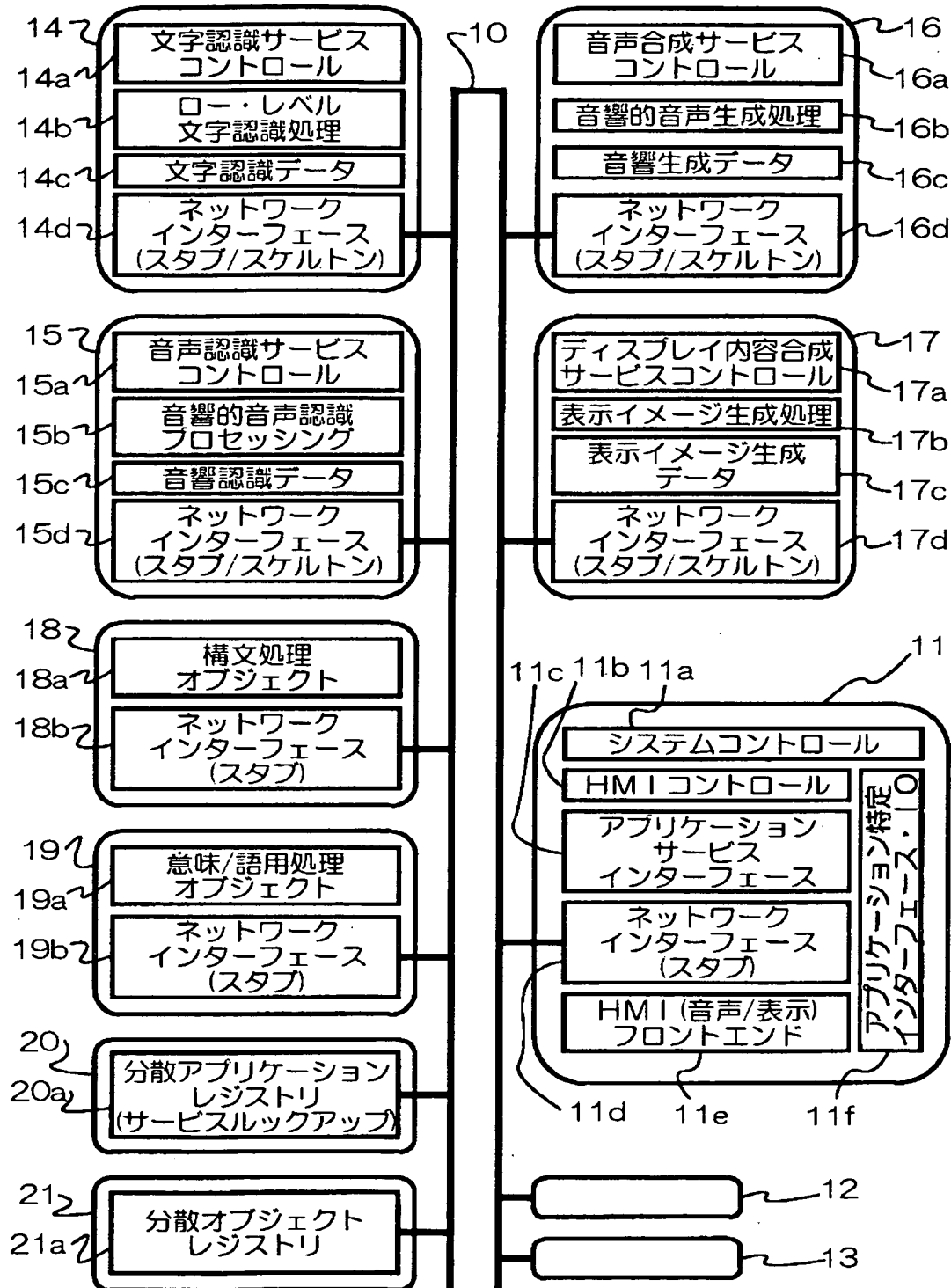
【図 10】



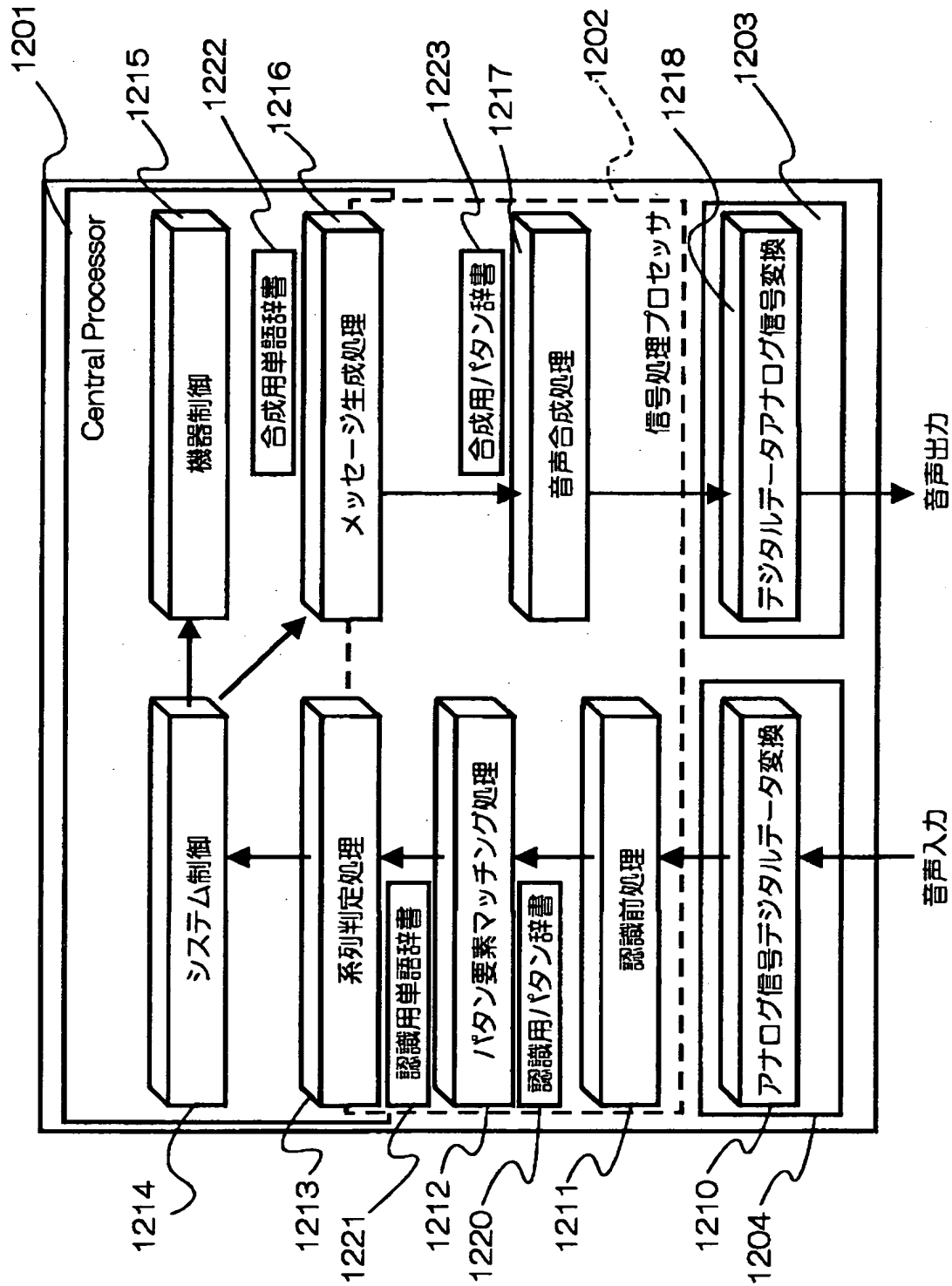
【図 11】



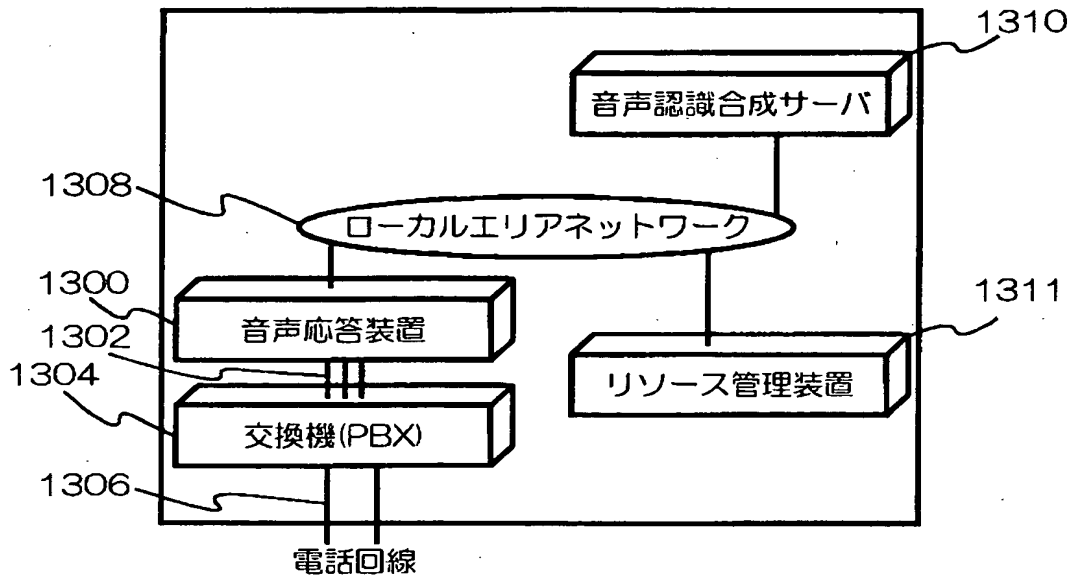
【図 12】



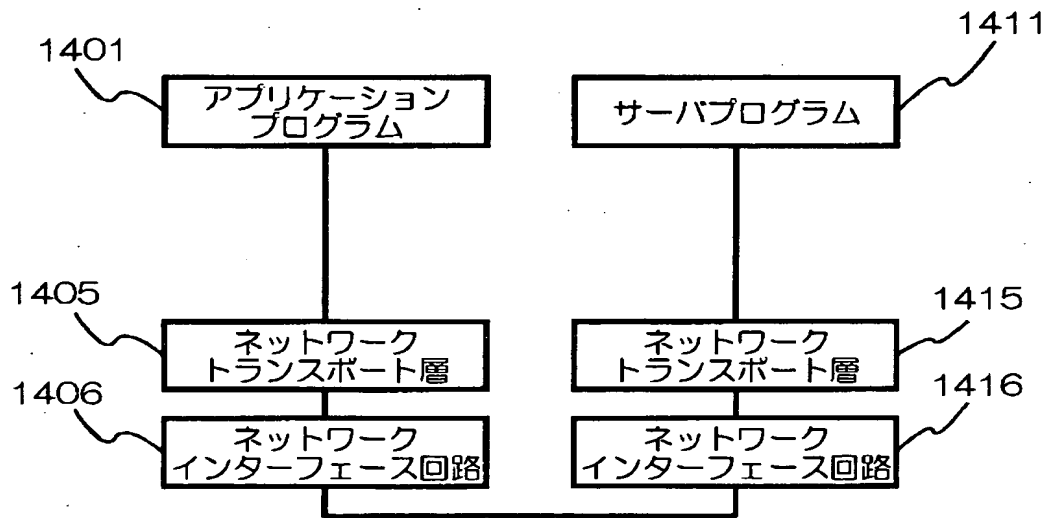
【図 13】



【図 1 4】



【図 1 5】



【書類名】 要約書

【要約】

【課題】 機器へのヒューマンマシンインタフェースの組み込み、実装にあたって、装置あたりのコストを低減したり、操作上の情報を機器間で共通化して異なる機器であっても同一の操作感が得られるようにしたりすることを目的とする。

【解決手段】 ヒューマンマシンインタフェースの機能を各々が分割して分散オブジェクトの形態で構成する各機能要素を、ネットワーク100のノード101～107上に複数配備し、各ノードが連携してデータ処理を行うことで、各ヒューマンマシンインタフェースの機能を実現するものであって、各ノードが人間との間のデータの入出力部であるアプリケーションノード101,102,103,105と、アプリケーションノードとネットワーク100上で接続されていて入出力されるデータを処理するサービスノード104と、両ノードの機能を備える複合ノード106,107のいずれかである。

【選択図】 図 1

認定・付加情報

特許出願の番号	特願2000-215062
受付番号	50000896045
書類名	特許願
担当官	高田 良彦 2319
作成日	平成12年 7月25日

<認定情報・付加情報>

【特許出願人】

【識別番号】	000004237
【住所又は居所】	東京都港区芝五丁目7番1号
【氏名又は名称】	日本電気株式会社

【代理人】

申請人

【識別番号】	100108578
【住所又は居所】	東京都新宿区高田馬場3丁目23番3号 ORビル 志賀国際特許事務所

【氏名又は名称】	高橋 詔男
----------	-------

【代理人】

【識別番号】	100064908
【住所又は居所】	東京都新宿区高田馬場3丁目23番3号 ORビル 志賀国際特許事務所

【氏名又は名称】	志賀 正武
----------	-------

【選任した代理人】

【識別番号】	100101465
【住所又は居所】	東京都新宿区高田馬場3丁目23番3号 ORビル 志賀国際特許事務所

【氏名又は名称】	青山 正和
----------	-------

【選任した代理人】

【識別番号】	100108453
【住所又は居所】	東京都新宿区高田馬場3丁目23番3号 ORビル 志賀国際特許事務所

【氏名又は名称】	村山 靖彦
----------	-------

出 願 人 履 歴 情 報

識別番号 [000004237]

1. 変更年月日 1990年 8月29日
[変更理由] 新規登録
住 所 東京都港区芝五丁目7番1号
氏 名 日本電気株式会社